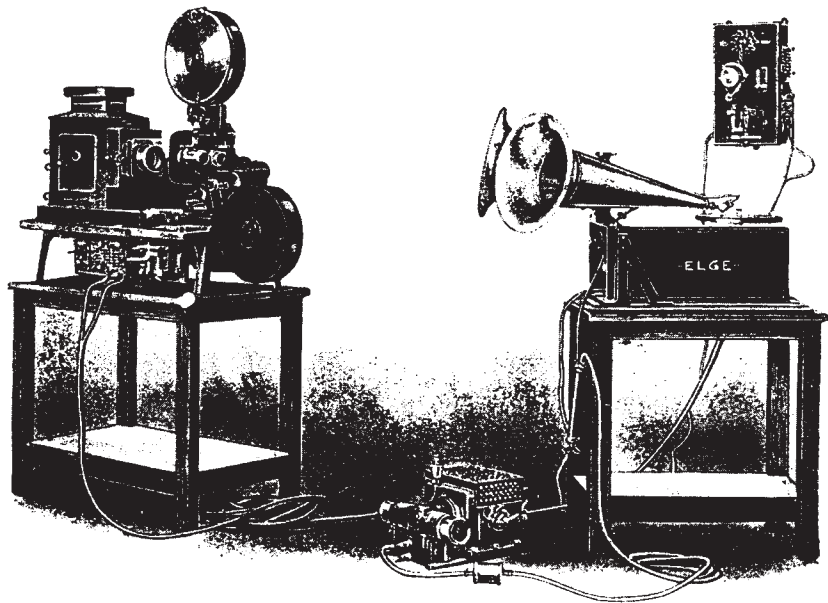


Alain Boillat

# DU BONIMENTEUR À LA VOIX-OVER



Ce livre en format pdf ne contient pas de DRM et peut donc être placé sans limitations sur vos ordinateurs, tablettes et téléphones.

Son usage est cependant limité aux cercles familial et amical.

Merci de nous contacter pour un usage plus étendu:  
[editions@antipodes.ch](mailto:editions@antipodes.ch)

# **DU BONIMENTEUR À LA VOIX-OVER**

**VOIX-ATTRACTION ET VOIX-NARRATION AU CINEMA**

## MÉDIAS ET HISTOIRE

Collection dirigée par Gianni Haver

Les médias occupent un rôle si fondamental à partir du XIX<sup>e</sup> siècle qu'ils deviennent des objets incontournables dans l'écriture de l'histoire contemporaine. Source privilégiée pour l'histoire culturelle et des mentalités, ils sont aussi au centre de nombreuses problématiques de l'histoire politique et sociale. La collection *Médias et Histoire* est ouverte à toutes les approches qui considèrent l'étude des représentations comme un moyen de comprendre la société, l'imaginaire collectif comme un élément important de l'exploration historique et l'analyse de la réception comme un questionnement et une confrontation nourrissant la réflexion en histoire.

## COMITE SCIENTIFIQUE

François Albera (Université de Lausanne), Hervé Dumont (directeur de la Cinémathèque suisse), Laurent Guido (Université de Lausanne), Vinzenz Hediger (Ruhr-Universität Bochum), Pierre-Emmanuel Jaques (Université de Lausanne).

## REMERCIEMENTS

Cet ouvrage, extrait d'une thèse qui a obtenu en octobre 2006 le Prix « Sciences humaines » de la Banque Cantonale Vaudoise, a été publié avec le soutien du Fonds des publications de l'Université de Lausanne, de la Fondation du 450<sup>e</sup> anniversaire de l'Université de Lausanne, de la Section d'histoire et esthétique du cinéma de l'Université de Lausanne, de la Société académique vaudoise et de la République et canton du Jura.



## MISE EN PAGE

Claudine Daulte, [cl.daulte@bluewin.ch](mailto:cl.daulte@bluewin.ch)

## CORRECTION

Claude Paré, Lausanne

## COUVERTURE

*Image et son. La revue du cinéma*, N° 215, mars 1968, p. 121.

© 2007, Éditions Antipodes

Case postale 100, 1000 Lausanne 7, Suisse

[www.antipodes.ch](http://www.antipodes.ch) – [editions@antipodes.ch](mailto:editions@antipodes.ch)

ISBN: 978-2-88901-994-6 [version PDF]

DOI: 10.32551/ANTIPODES.19946 [version PDF]

Alain Boillat

**DU BONIMENTEUR À LA VOIX-OVER**  
VOIX-ATTRACTION ET VOIX-NARRATION AU CINEMA

Préface de François Albera

Éditions Antipodes

## **MES REMERCIEMENTS**

A François Albera, directeur de thèse; à Jean-Michel Adam, Jean-Pierre Esquenazi et André Gaudreault, experts au colloque et à la soutenance; à Stefania Maffei, Laurent Guido, Olivier Liniger et Vincent Verselle pour leur relecture; à Édouard Arnoldy et Martin Barnier, avec qui j'ai eu d'enrichissants échanges.

# PREFACE

FRANCOIS ALBERA

«La voix continuait seule,  
prenait soudain cet accent, cette sonorité particulière,  
creuse, comme si elle se séparait d'elle...  
comme si voix et personne faisaient deux,  
chacune se mettant à vivre d'une vie indépendante, autonome,  
la première courant pour ainsi dire sur sa lancée.»  
(Claude Simon, *Le Vent*)

En 1946, Raymond Queneau donna à la revue *Labyrinthe* un petit texte démontant «le mythe du documentaire» où il commence par évoquer les séances de cinéma auxquelles, tous les dimanches de la guerre de 1914-1918, il assistait avec son père, dans la même salle du Havre, au second rang. Au troisième, juste derrière eux, une dame et son fils occupaient également toujours les mêmes fauteuils: «C'était une veuve bourgeoise et qui avait un fils idiot...»

«Le spectacle comprenait à cette époque un film de longueur moyenne, un comique, puis le grand film et enfin un second comique. La veuve et l'idiot accompagnaient le spectacle de leurs commentaires; ceux de l'idiot se réduisaient à grogner de temps à autre (Pathé frères, Pathé frères); ceux de la veuve présentaient par contre un intérêt certain. Cette dame, en effet, croyait que l'on projetait le même film tout au long de la séance, et elle essayait de retrouver les mêmes acteurs dans les deux comédies et les deux tragédies et d'y poursuivre la même intrigue. Son fils, l'idiot, tentait de la rappeler à la réalité, mais elle s'y refusait et revenait le dimanche suivant tout émerveillée.»

Vrai ou sollicité, ce souvenir met en place tout un dispositif dont le présent ouvrage d'Alain Boillat examine les composants dans le détail, analyse le fonctionnement, déplie les conséquences narratologiques, esthétiques, psychologiques, anthropologiques. Le « discours intérieur » du spectateur – dont parlait Boris Eikhenbaum en 1927 comme un élément structural de la production de sens au cinéma – est ici proféré; dès lors il rencontre le commentaire du bonimenteur du cinéma parlé, agent extérieur au film visuel mais participant du dispositif cinématographique, chargé de rendre intelligibles les films « des premiers temps » aux premiers spectateurs. La veuve est spectatrice, elle est donc un bonimenteur « amateur », phénomène moins rare qu'on ne le croit, qui perdure et que l'approche pragmatique aura un jour à construire comme une catégorie. Enfin, il s'apparente au commentaire *off* – mais on comprendra, à lire Boillat, qu'il faut désormais dire *over* – ou à telle ou telle voix « acousmatique » dont le cinéma parlant est coutumier. Ces trois « voix » participent à des degrés divers à ce que Boillat appelle le phénomène de la « déliaison vocale », déliaison entre la voix et l'image. Tous trois représentent une instance de verbalisation qui va à la rencontre de la suite d'images projetées et lui dispute la fonction référentielle. Boillat développe cette question sous le nom de *coréférence* – simultanée, rétrograde ou différée – et sur fond de comparaison intersémiotique.

Il est vrai que la dame « coopère » moins qu'elle ne s'approprie le film et lui fait violence – on pourrait y revenir – alors que le bonimenteur se met à son service, l'explicite, et que la voix *over* participe du film. Boillat les distingue nettement en utilisant le critérium de la provenance humaine (corps vivant) ou machinique (évicton de la dimension physique) de la voix et de la situation du locuteur. Mais il relève aussi les déplacements qui s'opèrent de l'un à l'autre (le retour du bonimenteur dans le cinéma parlant avec *Le Roman d'un tricheur* de Guitry).

Il n'empêche. Il y a disjonction à la base du film : que le commentaire ou le dialogue (ou le son en général) soient dus à un bonimenteur présent dans la salle ou caché derrière l'écran, que ces sons soient enregistrés par une autre machine que celle de prise de vues (ayant ses particularités de captation-altération, de fixation, de restitution) et diffusés parallèlement à la projection (plus ou moins synchronisés – disques), enfin qu'ils soient disposés sur la



pellicule (fixés, inscrits) mais sur une partie distincte de la partie image (et qui demeure quoi qu'il en soit décalée en raison de l'emplacement du lecteur optique ou magnétique du son sur l'appareil de projection). Cette disjonction que peut assumer la «voix-attraction», peut être suturée – dans la recherche d'un effet de réalité (synchronisme labial – la «voix-action»), par volonté de cohérence textuelle (le commentaire – «voix narration») – mais on peut aussi en écarter les pôles. Eisenstein a pu dire qu'il s'agissait de «monter» la parole et l'image comme on monte des images fragmentaires en obtenant des effets de sens qui sont fonction de l'éloignement des espaces de référence de chacune (attraction). La voix *over*, affirme d'ailleurs Boillat, actualise le discours élaboré par le montage du film. On peut entendre cette proposition extensivement.

À la lecture de ce livre, on ne peut manquer d'être convaincu que la question de la dissociation image/son (déliation, désancrage, distance, extériorité) est la grande question du cinéma. Car c'est bien le dispositif cinéma qui la rend possible: «il y a nécessairement, au cinéma, une *déliation fondamentale* des sons par rapport à l'image qui est constitutive du support de la représentation». Et que, selon l'intuition ou la conviction de Boillat, il est juste de penser que c'est en partant du «cas» de la voix *over* qu'on pénétrera le mieux dans ce problème pour en faire l'instrument d'une réélaboration générale de la production du sens au cinéma.

La voix *over* se distingue en effet de la voix *in* (émanant d'un locuteur visible à l'image, appartenant évidemment à la diégèse – désignation) et de la voix *off* (associée par le spectateur à un locuteur diégétique hors-champ – inférence), par le fait qu'elle renvoie à une source verbale qui est un énonciateur absent du «monde du film». Cette voix renvoie donc à un sujet humain dont on n'a pas d'image – qui n'a donc pas de corps (bien qu'elle ait cependant «du corps», une matière, une «étoffe», un «grain»). Il y a là un paradoxe à la base de la dialectique qu'examine l'auteur entre cette «intrusion d'un facteur humain» (voix = personne) et cette dimension machinique (impersonnelle). En effet, si «le mouvement anime les choses (...), seule la voix «anime» les êtres», écrit-il en une formule qui pourrait faire date: c'est un phénomène technique d'enregistrement-restitution via une série d'opérations qui captent certes mais aussi bien transforment, altèrent, qui insufflent une «vie» imaginaire à des ombres projetées, planes, grises

ou colorées ; technicité, mécanique « humaines, trop humaines » en quelque sorte !

Ce phénomène « de base » mobilise des questions qui précèdent et excèdent le cinéma (Boillat est remonté dans sa thèse de doctorat aux « machines parlantes » du XVIII<sup>e</sup> siècle, aux automates, et est passé par les extrapolations de Villiers de L'Isle-Adam, Jules Verne, Charles Cros et par les recherches d'Edison) ou l'accompagnent (Cendrars) et qui s'inscrivent dans une perspective anthropologique, du type de celle qu'avait ouverte Edgar Morin sur un autre plan avec son « homme imaginaire » – qui restait largement du côté de la seule image. En se situant cette fois du côté du corps du spectateur, comme le Ferdinand de *Pierrot le fou* (« une machine pour voir qui s'appelle les yeux, pour entendre les oreilles, pour parler la bouche... ») et de ce fait du côté du corps du film. En saisissant en effet combien, en raison du fait que la voix ne soit pas assignable à un corps – ni présent dans la salle, ni représenté sur l'écran – on est conduit à faire advenir le « corps du film » que le spectateur, « privé » lui-même de corps par sa situation dans le dispositif, est appelé à « investir ». La voix *over* dessine « un lieu à habiter, un non-corps à revêtir qui est le corps du film » (selon une formule de Mary Ann Doane).

Après une première « vague » dans le film noir américain des années 40 qui fut étonnamment audacieux (un mort peut se faire le narrateur de ses dernières heures), puis chez Welles, Cocteau et Bresson notamment, une certaine « modernité » cinématographique a fait fond sur cette disjonction (Resnais, Godard) et cet avènement du corps filmique (Debord, Straub) via cette opération technique d'association, de montage, pour dissocier, opposer, plaquer, étrangéfier, éloigner la référence du texte iconique de celle du texte verbal au point de différer toute unité diégétique (*Son nom de Venise dans Calcutta désert* de Marguerite Duras) ou de la rendre fondamentalement instable (*L'Homme qui ment* de Robbe-Grillet).

Il vaut la peine de rappeler ici que le présent ouvrage procède de la *réduction* d'une thèse universitaire de grande taille dont le point de départ était (relativement) restreint puisqu'il s'agissait d'analyser la question de la voix *over* dans les films d'Alain Resnais (l'étude de cas portant sur *Hiroshima mon amour* en témoignage). Or le dessein théorique de l'auteur – dont la formation ressortit à la fois à la

linguistique française (narratologie, énonciation) et à l'histoire et esthétique du cinéma (avec un goût prononcé pour les travaux sémiologiques) – a rencontré deux nécessités dans le cours de sa recherche : celle d'élargir son corpus à tout un ensemble de films autres que « l'œuvre-de-Resnais » et, au-delà même des films, de textes (*L'Ève future* est une des matrices du propos), et celle de se confronter à l'histoire du cinéma et à ses conséquences pour la théorie (aspect qui, jusqu'à Burch, Gaudreault, Gunning et quelques autres, était demeuré trop ignoré des chercheurs avec les importants effets qu'ils ne surent prévenir sur leurs modèles théoriques). Cette confrontation l'amena dès lors à élargir le champ de son investigation, d'abord taxinomique, jusqu'à cette dimension anthropologique dont on vient de parler. Il surmonte ainsi deux faiblesses de son stimulant premier ouvrage, *La Fiction au cinéma*, qui utilisait des exemples de films de manière fragmentaire et sans considération de leur ancrage historique.

Le renouvellement radical de l'investigation historique impulsée par la redécouverte du « cinéma des premiers temps » (longtemps laissé dans l'ambiguïté du « primitif », de la gestation) a permis cette ouverture qui n'en finit pas de bouleverser les idées reçues, les certitudes et les « intuitions » qui concernent le cinéma. On le sait, cette série de retournements est venue des chercheurs américains avant tout (québécois et états-unis), et, contrairement à plusieurs pays d'Europe (Italie, Pays-Bas...), elle peine à se faire jour en France – si l'on excepte le pionnier Noël Burch – où la tradition historiographique et théorique, sans doute d'avoir été parmi les plus riches, en est venue à former un obstacle à cette compréhension.

Le « bonimenteur » qui a émergé dans ce cadre est ainsi *l'œuf de Colomb* du « cinéma des premiers temps ». Aux interrogations qui se firent alambiquées à force de demeurer encloses dans une perspective mono-iconique et narratologique – donnant lieu à l'expression discutable de « style non-continu » – ne cessent dès lors de se substituer des réponses « évidentes » : le spectateur imaginaire du théoricien des années 1970-1980 était supposé ne pas comprendre ceci ou cela, s'étonner de telle autre disposition du « langage primitif » ou assister à la mise en place « des codes narratifs », alors que la prise en compte de la voix explicative ou narrative, du commentaire, de la description verbale ou de la sonorisation tire ce spectateur de ses supposés handicaps et les faiseurs de

films de leurs balbutiements. Boillat est un des premiers à tirer toutes les conséquences de cette redécouverte qu'il relie à la théorie, révolutionnaire elle aussi, du « discours intérieur » de Boris Eikhenbaum, inscrivant de la sorte la question de la voix *over* dans une généalogie qui part de l'oralité, de la performance – selon Paul Zumthor – et aboutit à la voix enregistrée. La distinction cinéma *parlé*/cinéma *parlant* est particulièrement suggestive et riche avec les caractères respectifs de l'un et l'autre : oral, labile *vs.* inscrit, fixe en particulier. L'une des conséquences de ce parcours évolutif est de mettre en exergue la nature *représentationnelle* des sons, dimension souvent occultée en raison d'une apparente « naturalité » du phénomène sonore au cinéma (on a pu prétendre qu'il n'y avait pas de différence entre son naturel et son enregistré, contrairement au rapport de l'image à son objet qui relève d'un codage).

Si l'histoire est abordée d'emblée comme un substrat riche de suggestions théoriques – la théorie de l'énonciation est reprise dans une perspective qui s'annonce comme « matérialiste » car elle refuse l'angélisme bien souvent de mise qui néglige les conditions concrètes, matérielles, physiques de la projection au sein d'un dispositif donné – l'esthétique du cinéma n'est en rien « sacrifiée » par ces perspectives historiques et théoriques. On se rappelle que l'auteur voulait à l'origine s'expliquer un aspect du « style » de Resnais ; à l'arrivée il a envisagé tout un corpus de films où ces considérations concernant la voix-*over* sont problématisées : ceux de René Clair sont sans doute les moins attendus de nos jours, tant ce cinéaste, majeur dans les années 1930 à 1950 en France, en Italie, en URSS, voire aux États-Unis, a disparu des références. Ce n'est pas le lieu d'en discuter (les discredits *critiques* sont, dans tous les arts, un de ces phénomènes qui rythment les hiérarchies, les redécouvertes ou les occultations), mais il faut souligner qu'Alain Boillat, saisissant un trait continu du travail de Clair, tire un fil d'importance : celui de la réflexion « en acte » qu'a menée ce cinéaste sur les rapports son-image et verbal-iconique, d'*Entr'acte* à *Tout l'or du monde*, et qu'on ne peut comparer qu'à celle qu'a entreprise et poursuit un Godard de nos jours – dans un tout autre espace, cela va sans dire<sup>1</sup>. De même Guitry, Duras ou

1. Cette réflexion sur la « voix déliée » chez René Clair sera proposée par l'auteur dans une publication ultérieure.

Resnais justement – sans parler du massif du cinéma expérimental (Michael Snow en particulier). Même si les analyses de films participent intimement du développement théorique, elles ont leur validité propre.

L'ouvrage de Boillat ne laisse rien au hasard : dans les différents champs à l'intersection desquels il dessine la place de son objet – sémiologie, histoire, psychologie – on peinera sans doute à repérer des omissions ou des lacunes dans les références qu'il cite et qu'il travaille. C'est que la clarté des analyses et leur systématité sont inséparables d'une discussion pied à pied avec ses prédécesseurs dans les questions traitées et le repérage de leurs limites. On pourra trouver parfois le jeune chercheur « insolent » avec ces aînés – qui ont noms Barthes, Eco, Metz, Genette, Hamburger, Milner, Schaeffer ou encore Mitry, Burch, Gaudreault, Chateau, Chion, Bradigan, Châteauvert... – mais on doit avant tout saluer l'énergie mise à renouveler des réflexions – souvent présentées par leurs auteurs comme des essais, des recherches – qui se sont trouvées figées et instrumentalisées dans les manuels et les dictionnaires (genre « littéraire » trop prisé par les éditeurs tandis que la recherche est tenue en lisière). La « mécanique » narratologique à l'usage des élèves du Secondaire (voire du Primaire : le modèle actantiel de Greimas !) a fait perdre toute productivité aux explorations et questionnements des auteurs ci-dessus et de bien d'autres dont il s'agit désormais de « reconnaître » les schémas et les classifications canonisés par des pédagogues fonctionnalistes.

Or, ces vigoureuses discussions chez Boillat n'ont rien d'académique, elles comportent des enjeux de taille sur la nature des études « sur la signification au cinéma » – pour reprendre un titre célèbre. Ainsi la question de l'énonciation au cinéma arrachée à une dérive anthropoïde (elle est pour lui un « simulacre impersonnel produit par une machine ») et confrontée à la théorie des « mondes possibles », celle de la coréférence verbo-iconique, l'intersémiotité – dont il écrit que « l'hégémonie de la narratologie genettienne » éloigna les théoriciens du cinéma – celle de la construction de la diégèse filmique, l'inscription de la sémiologie du cinéma dans une pragmatique. Sur tous ces points cardinaux en théorie du cinéma, l'auteur fait des mises au point magistrales et développe des propositions qui redonnent une vigueur et une actualité parfois inattendues à des démarches qu'on croyait appartenir à une époque un peu « oubliée » de la recherche, abolies par

d'autres modes intellectuelles (les philosophes et les esthéticiens ayant supplanté les linguistes et les sémiologues). Boillat se fait ainsi l'éveilleur de toute une série de questions élaborées par la sémiologie et qui se révèlent des instruments toujours nécessaires. Son mérite n'est pas mince à cet égard.

Par ailleurs, soulignons la «gourmandise» et «l'appétit» de l'auteur – presque de la boulimie – dans sa capacité à accueillir tout ce qui peut abonder à son entreprise: on verra les références classiques, légitimes, qu'on a déjà évoquées, jouxter des travaux de doctorants encore inédits, des traités techniques, des récits de témoins, des articles de critique, le recours à une vaste littérature où rien n'est négligé, les théories, fussent-elles datées, valant pour leur indicialité d'une manière de poser un problème à un moment donné autant que pour leur propos explicite. Qu'un théoricien se livre à une enquête aussi exhaustive – dans trois aires linguistiques au moins de surcroît – est un fait suffisamment rare pour être souligné. Et surtout, intégrer l'imaginaire du cinéma – via les textes critiques, esthétiques ou plus prosaïquement narratifs – à l'examen de l'imaginaire au cinéma, c'est précisément adopter un point de vue qui part des faits pour s'élever aux concepts et non, à l'inverse, qui veut plier les premiers aux seconds!

Tout travail d'investigation théorique qui ne s'en tient pas à des certitudes et ne se borne pas à la vulgarisation a pour pierre de touche sa capacité à voir son champ d'application élargi ou ses «modèles» exportés ailleurs. À l'évidence, c'est ce qui se passe ici, et le lecteur qui se jette dans le sillage du «Boillat», suit ses développements et éprouve les débats qu'il ouvre, vérifie leur pertinence en refaisant le chemin des exemples allégués, aura, à la fin de chaque chapitre mais souvent de chaque sous-partie, l'envie de rechercher ses propres exemples, d'appliquer et d'extrapoler, de faire marcher à son profit cette machine de précision – quand l'auteur lui a laissé quelque chose qu'il n'ait déjà passé au crible! Le jeu est infini et il engage chacun pour lui-même, mais songeons à l'apport qu'aurait, dans la perspective ici dessinée, la prise en compte de l'œuvre si singulière d'Alexandre Sokourov dont le travail sur le son et les voix, toujours décalées par un renoncement de principe au son direct, flottantes, jouant de tous les registres de leur matière (accent, souffle, borborygme, bredouillement, éructation, bégaiement, aphasie), le recours extensif à la voix *over* dans nombre de ses films (de *Confession* à *Voix spirituelles*, de *Humble vie*

à *Élégie de la traversée*) – est certainement l'un des plus inventifs du cinéma contemporain.

« Qui parle ? » en effet dans le film, et d'ailleurs où commence le « film » ? Ces questions ouvertes ici se trouvent plus que jamais relancées par la production contemporaine, et un ensemble plus vaste où les jeux vidéo jouxtent le CD-Rom, les sites web et de nouvelles « textualités » audiovisuelles, l'auteur s'en avise dans sa conclusion.

Les séances des « premiers temps » ont pu être envisagées comme unité textuelle de base (on va à la séance – plurielle, hétérogène [voir Queneau] – non voir *un* film) mais l'institutionnalisation, l'instauration du long-métrage, l'abandon de pratiques spectaculaires extrinsèques au film même (bonimenteur, attractions dansées, chantées, etc.) ne doit pas trop vite faire conclure à une redéfinition plus étroite – qui s'est pourtant imposée par forclusion de tout ce qui paraissait attenter à la sacro-sainte « spécificité ». L'actuelle pluralisation des pratiques de visionnement des films comme la diversité des modalités de spectacles demandent sans doute que l'on admette des définitions variables des notions de « film » et même de « texte filmique ». Le schéma que propose l'auteur à propos du bonimenteur au sein du dispositif du cinéma parlé montre bien que la représentation visuelle (projection de vues animées) se retrouve, avec la représentation sonore assurée par le bonimenteur et la représentation visuelle *live* (quand il y en a une), au sein d'une entité englobante appelée « méga-représentation audio-visuelle » qu'on pourrait envisager comme définissant le « film » de l'époque.

Boillat écrit que l'étude de la *coréférence* permet de comprendre comment la « voix de la machine » s'avère capable de construire un monde, à l'instar d'un conteur qui nous livrerait un récit oral, mais « en proposant (et en imposant) un réglage strict des processus sémantiques ». *Proposant* et *imposant*, ces deux termes importent manifestement beaucoup à l'auteur, qui tient à analyser un « fonctionnement textuel » que, certes, « active » le spectateur, mais qu'il ne saurait transformer... On revient à la dame de Queneau – une voix dans le dos, dans le noir – qui construit, à partir des différents « films » (*textes distincts*), un seul récit par la grâce de l'oralité, un monde possible donc, sans se borner à « activer » la structure textuelle – selon le schéma postulé par Eco. Elle prend une place à elle que le dispositif lui permet d'occuper mais qu'il n'a pas

«prévue». Oralité *live*, improvisation, performance, voix humaine qui affronte la bande machinique du film, elle parvient à changer le sens du message émis. Queneau, que passionnent les stratégies de détournement et les combinatoires aléatoires, ne s'y trompe pas qui explique dans la suite de son article qu'il a souvent appliqué la méthode de cette dame dans ses lectures ou ses visions de films.

C'est peut-être là que s'arrête Boillat et que commence une autre histoire, celle des spectateurs et de leurs pratiques de «spectature» – le mot a été proposé par Martin Lefebvre dans un sens interprétatif – qui sont des appropriations. La pratique des surréalistes jouant de l'effraction, celle des situationnistes pratiquant le détournement n'ont guère été prises en compte par les théoriciens, mais pas plus les séances «a-typiques» dont a pu faire état Bazin après un voyage en Algérie et une série de projections itinérantes parmi des populations peu accoutumées au cinéma, ou Jean Rouch quand il vint montrer aux Dogons les films qu'il avait faits de leurs rituels, ou encore plus prosaïquement les projections de films de famille. Or les technologies actuelles en rendent l'exercice de plus en plus fréquent par un biais évidemment tout autre: on propose aujourd'hui sur un site internet des films de Bollywood détournés par sous-titrage. L'exercice – inauguré il y a déjà quelques décennies par les situationnistes avec des films chinois – est limité: il tient à l'ignorance de la langue parlée par les protagonistes du film. Mais il est l'indice d'une complexité jusqu'ici ignorée dans les modalités de réception et d'appropriation des films qui n'a, de fait, jamais cessé: bandes-son de films diffusées à la radio, films transcrits en rouleaux d'images fixes pour projections lanternistes, en romans-photos ou en romans, remakes, remontages, etc. Autant de pratiques qui donnent une fonction prééminente à la voix-over, à une «oralité» dont Paul Zumthor, l'auteur le rappelle, annonçait le retour.



## PROLOGUE

Généralement, le cinéma est prioritairement associé à la présence d'images. Or le défilement photogrammatique ne constitue qu'une seule des cinq « matières de l'expression » cinématographique, aux côtés de la voix, des bruits, des mentions écrites et de la musique. Au sein de la pluralité sémiotique du médium, la voix me paraît être une composante décisive, et cela tout au long de l'histoire du cinéma: non seulement elle résonnait déjà dans les salles à l'époque improprement nommée « muette », inscrivant les projections dans l'héritage de pratiques antérieures, et s'est imposée de façon extrêmement rapide sous une forme fixée à partir de la fin des années 20, mais elle constitue à divers égards un facteur essentiel de l'organisation discursive des films parlants. Cette fonction est d'autant plus patente lorsque la voix apparaît sous une forme *over* où, détachée du monde du film, elle commente, raconte, nous fait pénétrer dans un environnement ou dans la conscience d'un personnage. Ce sont ces différents aspects dont je projette de traiter dans cet ouvrage.

### LA VOIX ET L'IMAGE, OU L'HUMAIN DANS LA MACHINE

Au cinéma, la voix constitue un facteur déterminant à plusieurs égards: en tant que véhicule du langage verbal, elle contribue à la production du sens et à la mise en place de la diégèse filmique; résultat d'une profération, elle postule l'existence d'un énonciateur qui entretient différents types de liens avec l'image; matière vocale, elle soumet, à des degrés variables, la transmission du message audiovisuel au régime de l'oralité. Ce dernier point présente l'avantage méthodologique d'exiger un cadre conceptuel intermédiaire, et de susciter une réflexion sur l'intervention d'un facteur

emblématique de l'humain dans un moyen d'expression caractérisé par sa nature mécanique. On verra que la voix des films parlants, bien qu'appartenant en propre à un individu, est en outre déterminée par la matérialité technique de son enregistrement et de sa restitution. Elle représente par conséquent l'agent d'une rencontre entre l'humain et la machine dont l'aspect conflictuel me paraît productif d'un point de vue théorique. Il s'agira de prendre la mesure des implications de l'oralité et, inversement, de la rupture épistémologique introduite par la généralisation de l'inscription des sons. Ainsi, sur la base d'une étude des manifestations vocales dans le contexte de l'audiovisuel, je serai amené à proposer de nouveaux modèles d'analyse filmique, et à rediscuter certains acquis théoriques jusqu'ici trop informés par le primat accordé à la dimension visuelle.

Précisons que le titre du présent ouvrage n'entend connoter aucune conception évolutionniste. En effet, il ne s'agit pas tant de suivre une trajectoire qui partirait du bonimenteur du cinéma des premiers temps pour aboutir progressivement aux films à voix-over que de mener une analyse comparée de la voix vive et de la voix fixée lorsqu'elles sont couplées à des projections d'images animées. C'est pourquoi les spécificités du boniment ou de la voix-over seront également traitées en dehors des chapitres spécifiquement consacrés à ces deux types de manifestations vocales (voir les occurrences de ces notions dans l'index en fin de volume), ce qui permet à mon sens de mieux circonscrire les enjeux induits par chacune d'elles. En outre, les deux régimes dominants que je nommerai *voix-attraction* et *voix-narration* résultent d'une conception graduelle et dialectique: les phénomènes ne sont pas envisagés comme se supplantant diachroniquement, mais comme intervenant de façon concurrente et à des degrés variables.

Une rupture survient évidemment à l'époque de la généralisation du cinéma parlant, lorsque les voix enregistrées évincent définitivement, au sein des pratiques dominantes en Occident, l'exercice de la voix vive qui incombait à un locuteur présent dans l'espace de la salle. Sur la base de discours contemporains des spectacles concernés, on comparera avec profit l'époque du boniment (1895-1914) avec la phase de standardisation du parlant (environ 1927-1935) afin de saisir les rôles respectivement endossés par l'humain et la machine. La voix fixée ne s'est d'ailleurs imposée ni subitement – les tentatives remarquées de sonorisation sont

presque aussi anciennes que les premières bandes cinématographiques – ni sous une forme monolithique. La période que Noël Burch a qualifiée d'« interrègne », qui recouvre les années 1926 à 1934, se caractérise par une forte diversification des pratiques<sup>2</sup>. J'interrogerai l'intuition du théoricien en accordant une attention toute particulière à cette époque de l'histoire du médium, notamment en discutant les liens qui unissaient les spectacles de vaudeville et le cinéma (par exemple à travers la pratique parfois mixte du prologue) et en étudiant dans le détail, au chapitre III, *Le Roman d'un tricheur* (Sacha Guitry, 1936). Ce film se présente à mon sens de façon emblématique comme l'une des dernières rémanences de l'interrègne. Parent à divers titres de pratiques appartenant à ce que j'appellerai le « cinéma » *parlé* (c'est-à-dire le « muet » *sonorisé live*), il témoigne en effet d'une exceptionnelle hybridité au niveau de l'utilisation de la voix. D'autres films ultérieurs font également un emploi productif de l'oralité, comme *Lola Montès* de Max Ophüls (voir IV.8) ou, en ce qui concerne plus spécifiquement la voix-over, *Hiroshima mon amour* d'Alain Resnais (voir le chapitre VII). On verra que certains usages de la voix au cinéma peuvent avoir des conséquences esthétiques notables, voire se profiler comme des innovations majeures.

### AXES METHODOLOGIQUES

Afin de ne pas sous-estimer les implications de la fixation des sons, il me paraît indispensable de mettre l'accent sur la matérialité technologique du cinéma parlant. À cette fin, je suivrai la proposition de James Lastra qui consiste à distinguer, en remontant les étapes de la genèse des sons filmiques, trois niveaux élaborés sur le modèle proposé pour l'image par le filmologue Étienne Souriau : le niveau *prophonographique* concerne ce qui s'est trouvé à proximité de l'appareil au moment de l'enregistrement ; le niveau *phonographique* s'applique au support de la fixation des sons ; enfin, le *postphonographique* a trait à toutes les opérations associées au montage<sup>3</sup>. On verra que cette dimension matérielle n'est pas sans incidences sur les questions d'énonciation filmique.

2. Noël Burch, « Du muet, le parlant. Réflexions cursives sur un interrègne », in Christian Belaygue (dir.), *Le Passage du muet au parlant*, Toulouse : Cinémathèque de Toulouse, 1988.

3. James Lastra, *Sound Technology and The American Cinema. Perception, Representation, Modernity*, New York ; Chichester : Columbia University Press, 2000.

Par ailleurs, je ne m'en tiendrai pas à la seule représentation audiovisuelle – dont le locuteur *live* semble a priori exclu – mais tenterai dans les deux premiers chapitres d'intégrer systématiquement l'origine vocale au sein d'un *dispositif*. Cette approche me permet de poursuivre le travail de Rick Altman qui, au début des années 90, renouvelait passablement l'étude des sons au cinéma en ouvrant de nombreuses perspectives dont la productivité est, il me semble, loin d'avoir été épuisée aujourd'hui. À partir d'un cadre théorique qui repose sur la notion de «cinéma comme événement» («cinema as event»), Altman postule de centrer l'analyse sur le contexte de production et de réception du film<sup>4</sup>. Plusieurs parties de mon étude viseront à développer certains points mentionnés dans le texte programmatique d'Altman. Je prendrai notamment en compte le médium au niveau «des ressources matérielles qu'il engage» et du large spectre de discours qu'il suscite; je m'interrogerai sur l'intersection de différentes sphères culturelles et sur la notion de «performance», que je spécifierai dans le cas de la parole en recourant aux thèses de Paul Zumthor. Altman situait toutefois son approche aux antipodes des considérations sur le «film comme texte», conception héritée du structuralisme dont il se servait comme repoussoir pour des raisons tout à fait compréhensibles à l'époque. Je proposerai quant à moi de réintroduire la composante «textuelle», non pas comme objet ultime de l'analyse, mais comme l'un de ses niveaux, d'ailleurs fortement déterminé «de l'extérieur» par différents champs culturels. La perspective du «cinéma comme événement» ne doit pas, à son tour, chasser toute considération sur le texte filmique, sous peine de faire resurgir le caractère limitatif qu'elle s'est évertuée à combattre.

Dans le champ des études sur le cinéma, les travaux qui ont conduit à un élargissement salutaire du domaine de recherche portent précisément sur la dimension sonore: à l'époque du cinéma «muet», les pratiques d'accompagnement des films présentaient une hétérogénéité et une labilité qui obligent l'historien à sortir de l'objet «film». Les sons n'existent que *diffusés* dans la salle, ce qui appelle une véritable diffraction des centres d'investigation. Rick Altman note à ce propos: «En tant que produit matériel, le cinéma révèle rapidement l'emplacement et la nature de ses pistes-son, la

4. Rick Altman, «General Introduction: Cinema as Event», in Rick Altman (dir.), *Sound Theory, Sound Practice*, New York; London: Routledge, 1992.

technologie utilisée pour les créer, le dispositif nécessaire pour la reproduction, et la relation physique entre les haut-parleurs, les spectateurs et leur environnement physique.»<sup>5</sup> Dans cet écheveau d'interactions, je tenterai de faire intervenir la voix vive du bonimenteur et les caractéristiques matérielles de la voix phonographiée.

La prise en compte de la réception filmique s'appuiera non seulement sur la construction théorique d'un spectateur tel que le déterminent une institution, un dispositif et une représentation donnés, mais également sur l'analyse des discours. Qu'il s'agisse de témoignages du début du XX<sup>e</sup> siècle ou de théories cinématographiques récentes, tous les textes seront considérés comme des *faits de réception*. Je ne me contenterai donc pas d'emprunter certaines notions à des auteurs d'importance (Roger Odin, Gérard Genette, André Gaudreault, etc.), mais je discuterai leurs travaux en tant que «sources», dans la mesure où ces textes sont souvent aussi intéressants pour ce qu'ils trahissent de la conception d'une époque (ou propre à une approche spécifique) que pour leurs propositions théoriques proprement dites. Si l'analyse filmique constitue dans cet ouvrage un instrument méthodologique important, mes investigations se déploieront au-delà des productions cinématographiques afin de saisir la façon dont le public conçoit, à une période historique donnée, le rôle de la voix. De nombreux textes appartenant à diverses catégories discursives seront par conséquent interrogés au niveau de ce qu'ils révèlent de la spécificité de la voix au sein d'un dispositif audiovisuel.

Deux principes interdépendants guideront la conceptualisation des différents paramètres considérés dans ce travail, facilitant ainsi leur mise en relation : d'une part le *caractère graduel* des notions concernées (la fictionalité, l'homodiégéticité, les fonctions du langage selon Jakobson, l'intégration narrative, la musicalité de la voix, le marquage énonciatif, etc.), d'autre part le *primat de la dimension pragmatique*. En vertu de ce second postulat, le gradient qui définit chacun de ces paramètres sur une échelle comprise entre une présence minimale et un pôle maximal n'est pas considéré comme étant propre au texte filmique, mais résulte d'une activité inférentielle incombant à l'*audiospectateur* (c'est ainsi que j'appellerai le destinataire d'un message audiovisuel lorsque la

5. *Ibid.*, p. 6.

composante sonore intervient de façon décisive). C'est pourquoi j'élaborerai un modèle de l'énonciation filmique dans lequel le marquage énonciatif sera fonction de sa perceptibilité par le spectateur. On verra comment les films occultent la réalité machinique du dispositif en incitant l'audiospectateur à se figurer mentalement une origine humaine de l'ensemble de la représentation, en adéquation avec un désir profond d'enracinement anthropocentrique de la «voix narrative».

La voix-over est donc un procédé qui convient idéalement à l'assignation (que je qualifierai de «*fictive*») d'une source humaine (souvent assimilée à un être *fictionnel*). C'est pourquoi je ne peux souscrire à l'hypothèse fréquemment soutenue selon laquelle la voix-over serait intrinsèquement un facteur de «distanciation» au sens de Bertolt Brecht. En discutant la notion de «discours intérieur» proposée dans les années 20 par le Soviétique Boris Eikhenbaum, je montrerai au contraire combien la voix-over se met au service de procédures visant à assurer l'immersion du spectateur dans le monde du film. On constatera que même un film caractéristique d'une certaine «modernité» cinématographique comme *Hiroshima mon amour* exploite les vertus humanisantes et unifiantes de la voix-over. Bien que prédisposée à une fonction de distanciation en vertu de l'immunité dont elle dispose envers les lois diégétiques, la voix-over conduit, dans les pratiques dominantes, à l'effet inverse, notamment en se posant en tant qu'instance d'instauration de l'univers filmique. Si, comme le dit Althusser, Brecht «voulait mettre le spectateur à distance du spectacle, mais dans une situation telle qu'il fût incapable de le fuir, ou d'en simplement jouir»<sup>6</sup>, la voix-over constitue quant à elle, dans la plupart des films, un moyen d'assouvissement du désir de fiction. En effet, à l'exception de quelques rares narrateurs hétérodiégétiques qui affichent une position de recul, les voix-over du cinéma parlant de fiction sont généralement utilisées à des fins psychologisantes qui se situent aux antipodes de la distanciation telle que l'entendait le dramaturge allemand. La seule énonciation d'un texte *over* au «je» favorise l'émergence de la conscience d'un personnage *focal* dont l'«intérieurité» nous est offerte de façon privilégiée, tandis que la position surplombante des énonciateurs *over* confère à

6. Louis Althusser, «Le «Piccolo», Bertolazzi et Brecht (Notes sur un théâtre matérialiste)», in *Pour Marx*, Paris: Maspero, 1975 [1<sup>re</sup> édition 1962], p. 146.

l'audiospectateur un pouvoir ubiquitaire dont se nourrit également sa pulsion scopique.

### VOIX IN, OFF, OVER: QUESTIONS DE DELIAISON

L'analyse des sons au cinéma, du point de vue de la (non-)visualisation de leur source, a donné lieu à de multiples classifications. Il est assurément indispensable d'étudier ces phénomènes dans leurs fondements esthétiques et leur impact sur le spectateur. Néanmoins, dans une entreprise visant à élaborer une conception graduelle et polyvalente de l'ancrage du son dans la représentation visuelle, il me paraît préférable de limiter l'inflation taxinomique en conservant la tripartition traditionnelle des anglophones (*in/off/over*), tout en la nuancant grâce à la prise en compte d'une configuration audiovisuelle particulière qui permet d'instaurer un continuum entre ces trois catégories: la *déliaison* (vocale). Certes, la distinction entre ces types de voix ne permet pas d'affiner l'analyse au-delà de situations types, mais elle offre quelques critères qui facilitent l'appréhension d'un ensemble de pratiques si diversifiées qu'un étiquetage plus précis s'avérerait de toute manière inadéquat, ainsi que l'illustrent le modèle de Daniel Percheron – dont le schéma arborescent constitue l'une des premières formalisations de la question, mais aboutit, dans l'analyse, à d'étranges catégories comme le son «off-in»<sup>7</sup> – et celui de Serge Daney, alourdi par l'ajout de deux types (la «voix *through*» et la «voix *out*»)<sup>8</sup>. Si l'on s'en tient à une distinction tripartite, on obtient les cas de figure suivants: grâce au principe du «synchronisme» entre la voix et les mouvements labiaux d'un acteur, la voix *in* est perçue comme l'émanation d'un individu visible à l'image en tant que source; la voix *off* est associée par le spectateur à un locuteur situé dans l'univers diégétique, à proximité du «point d'écoute» supposé, mais hors-champ; la voix-over est quant à elle rapportée à un énonciateur absent, en tant que source verbale, du monde du film (ou de l'un de ses mondes). Contrairement à l'usage largement répandu en France de la dénomination «voix off», j'opte ici pour

7. Daniel Percheron, «Le son au cinéma dans ses rapports à l'image et à la diégèse», *Ça Cinéma*, 1<sup>re</sup> année, 2<sup>e</sup> état, octobre 1973, p. 82.

8. Serge Daney, «L'orgue et l'aspirateur (Bresson, le Diable, la voix *off* et quelques autres)», in *La Rampe. Cahier critique 1970-1982*, Paris: Éditions des Cahiers du cinéma; Gallimard, 1983, p. 144.

le qualificatif «*over*» en suivant la proposition de Jean Châteaupert dont l'ouvrage, paru en 1996, a contribué à populariser cette expression parmi les chercheurs francophones<sup>9</sup>. Le terme «voix-over» (que j'écrirai en un mot pour renforcer son statut de notion-clé) me paraît en effet plus adéquat, d'une part parce qu'il permet d'éviter la confusion avec le sens anglais du terme «*off*» qui signifie «hors-champ» (ce qui est particulièrement malvenu dans ce contexte), d'autre part parce qu'il connote une dimension essentielle de ce procédé: la voix plane, tel un esprit, *au-dessus* du monde diégétique d'une part, dans l'espace de la salle d'autre part. Elle engage l'humain sans le corps, mais non sans emporter, à travers ce que Roland Barthes a appelé son «grain», une part de l'humanisation et de la corporéité du locuteur<sup>10</sup>. Le grain de la voix, comme le dit Châteaupert, constitue «une mine d'informations sur l'individu»<sup>11</sup>, le spectateur associant à certaines caractéristiques vocales un physique donné, quand bien même celui-ci n'est jamais montré à l'écran. En dépit de ces connotations, la voix-over provoque toutefois une désincarnation qui résulte non seulement de l'absence de visualisation du locuteur diégétique, mais aussi de la technique phonographique elle-même qui permet, ainsi qu'Edison le décrivait dans un texte de mai-juin 1878 où il exposait les possibilités de son invention, une reproduction réalisée sans tenir compte de la présence, ni même de l'existence de la source originale<sup>12</sup>.

Sur le plan de l'analyse filmique, le statut *over* est avant tout défini par rapport à la diégèse puisqu'il s'applique à un son qui, bien que parvenant au spectateur, ne peut être entendu par les personnages du film – le caractère exceptionnel et réflexif du non-respect de ce principe proposé par un film comme *L'Incroyable destin d'Harold Crick (Stranger than Fiction)*, Marc Forster, 2006) ne fait que souligner à quel point cette norme s'est généralisée et a été intégrée par les spectateurs. Selon la typologie genettienne, le locuteur *over* est par conséquent «extradiégétique», bien qu'il faille à mon sens discuter le sens de ce qualificatif ainsi que l'ensemble du système dont il est issu (voir le chapitre VI), que Genette désigne

9. Jean Châteaupert, *Des mots à l'image. La Voix over au cinéma*, Québec; Paris: Nuit Blanche; Méridiens Klincksieck, 1996.

10. Roland Barthes, «Le grain de la voix», in *L'Obvie et l'obtus. Essais critiques III*, Paris: Seuil, 1982 [1<sup>re</sup> édition: 1972].

11. Jean Châteaupert, «Il faut trouver la voix», *Cinémas*, Vol. 3, N° 1, 1992, p. 67.

12. Cité par J. Lastra, *op. cit.*, p. 19.



significativement par le terme polysémique de «voix». Relativement à l'espace acoustique de la salle, on peut dire par contre que tous les sons du cinéma parlant sont *over*, puisque les mouvements de l'écran ne produisent pas eux-mêmes les vibrations sonores perçues: il y a nécessairement au cinéma une *déliasion fondamentale* des sons par rapport à l'image qui est constitutive du support de la représentation. Les sons étant restitués par des moyens mécaniques ou informatiques, leur source effective est absente de l'espace de leur réception. Les voix du cinéma parlant offrent en cela une représentation ambiguë – et par là même fascinante – de l'humain, puisqu'elles évincent paradoxalement l'individu comme présence tout en s'en faisant de façon parfois très personnalisée la manifestation phonique. Inversement, si l'on cesse de rapporter la voix à l'image, ces catégories perdent toute pertinence car, comme l'a noté Christian Metz, «on oublie que le son lui-même n'est jamais <off>: ou bien il est audible, ou bien il n'existe pas»<sup>13</sup>. La référence obligée à la diégèse sur laquelle repose la distinction entre *off* et *over* exige d'une part de s'interroger sur les phénomènes d'élaboration de l'univers du film, d'autre part d'intégrer des paramètres extérieurs à la représentation écranique proprement dite, comme les conditions mécaniques de la diffusion sonore ou les différences existant entre des sons restitués et des sons effectivement produits par un locuteur dans la salle. Les critiques formulées à l'encontre du modèle *in/off/over*, d'ailleurs habituellement issues de recherches dans lesquelles la voix occupe une place mineure, résultent du constat qu'il n'est pas possible de rendre compte de certains cas sur la base de ces critères. Or, une conception graduelle de ces notions abroge cet obstacle méthodologique.

À cet égard, le modèle de Michel Chion, qui représente ces trois catégories à l'aide d'une figure qualifiée de «tricerclé», me paraît intéressant car il postule une perméabilité entre les types. Fondant sa schématisation sur un clivage entre les sons dont la source n'est pas visualisée – sons qu'il qualifie d'«acousmatiques» – et ceux qui sont rapportés à un élément de l'image, Chion souligne que l'on peut certes «la subdiviser et [...] la remettre en cause indéfiniment», mais qu'elle demeure pertinente pour autant que l'on reste sensible à l'existence de zones frontières<sup>14</sup>. On prêtera dans

13. Christian Metz, «Le perçu et le nommé», in *Essais sémiotiques*, Paris: Klincksieck, 1977 [1<sup>re</sup> édition 1974], pp. 157-158.

14. Michel Chion, *Le Son au cinéma*, Paris: Éditions des Cahiers du cinéma, 1985, pp. 32-34.

cet ouvrage une attention toute particulière à de tels cas limites qui permettront de concevoir la voix déliée comme un phénomène graduel, et cela également parmi les diverses manifestations de la voix-over. Il est cependant évident que plusieurs zones du spectre couvert par les possibilités de la déliaison vocale ne peuvent être abordées pleinement dans une étude consacrée à la voix-over<sup>15</sup>. Il s'agirait en effet d'élargir en deçà le champ d'investigation à l'idéologie de la voix synchrone (parangon de l'anthropocentrisme) et aux différentes modalités d'apparition des voix *off* et *déliées*, dont les effets sur l'audiospectateur sont variables<sup>16</sup>. En considérant deux extrêmes (les voix *in* et *over*) dans un rapport dialectique autorisant l'examen de plusieurs *régimes vocaux*, j'espère toutefois offrir un cadre général valable pour l'analyse des différentes voix du cinéma. Cet ouvrage s'organise en effet autour de quatre types de manifestations vocales qui, combinés selon des degrés variables, recouvrent à mon sens la plupart des utilisations de la voix au cinéma.

En raison de l'omniprésence de la voix dans des pratiques spectaculaires très diverses, les usages vocaux que l'on observe dans les films ne sont souvent pas spécifiquement cinématographiques. Plusieurs aspects dont il sera question ont d'ailleurs plus généralement trait au théâtre, forme de spectacle qui a passablement influencé les thèses de Paul Zumthor sur lesquelles je m'appuierai. L'intérêt de l'analyse réside précisément dans de tels emprunts intermédiaires. Sur le plan historique, je m'attarderai notamment sur les spectacles «de type vaudeville» qui font intervenir la voix sur un mode que l'on retrouve dans certains films. L'une des filiations que nous serons amené à constater, aux contours historiques aujourd'hui encore flous, résulte de la mise en rapport des prologues *live* réalisés dans la salle avant la projection avec ceux qui furent intégrés aux films, puis avec les voix-over inaugurales. Le mouvement de l'attraction à l'intégration (narrative) que décrit

15. Le présent ouvrage reprend en effet partiellement les chapitres 1, 5 et 6 de la thèse de doctorat intitulée *Les Voix du cinéma. Boniment, voix synchrone, déliaison et voix-over* soutenue en mars 2006 à l'Université de Lausanne. Même si j'esquisserai ici quelques pistes pour traiter de la déliaison (notamment sur la base des notions de «mise en phase», de «plaquage», d'«inquiétante étrangeté», etc.), le présent ouvrage demeure centré, en ce qui concerne le cinéma parlant, sur des cas de non-visualisation du locuteur, alors que la déliaison demanderait de considérer également divers flottements de l'ancrage vocal.

16. Voir à cet égard mon article «*L'Ève future* et la série culturelle des «machines parlantes». Le statut singulier de la voix humaine au sein d'un dispositif audiovisuel», *Cinémas*, Vol. 17, N° 1, automne 2006.

ce parcours historique de la voix liminaire, reflet de l'opposition plus générale entre oralité et fixation, correspond à celui qui guidera l'organisation des chapitres V et VI.

Une conceptualisation générale des manifestations vocales comme celle que j'entends proposer ici est inédite dans le champ des théories du cinéma, même si plusieurs points ont déjà été traités de façon isolée et partielle (le dialogue, la voix-over, l'adresse au spectateur, etc.). On constate en effet que la présence de la parole orale a été passablement négligée dans les écrits théoriques sur le cinéma. Dans le domaine francophone, l'un des rares auteurs à avoir fait preuve de véritables préoccupations théoriques envers la question de la voix est Alain Masson dans *L'Image et la parole*, où il entreprend d'analyser dans sa singularité «l'avènement» du parlant<sup>17</sup>. Cet essai ouvre la voie à une appréhension des phénomènes d'interaction entre la voix et l'image qu'il me semble judicieux de développer en intégrant notamment les connaissances acquises entre-temps sur l'histoire du son au «cinéma». Les recherches ultérieures des historiens travaillant sur les «sons du muet» permettent en effet d'approfondir et de préciser (voire de corriger) certaines conclusions de Masson en ce qui concerne les tentatives de sonorisation des films qui précéderent d'une vingtaine d'années la généralisation du parlant. La publication, dans la même année 2004, des résultats des recherches de Rick Altman, de Giusy Pisano et d'Édouard Arnoldy est révélatrice de l'intérêt rencontré récemment par cette question dans le champ académique<sup>18</sup>, auquel n'appartient indubitablement pas l'ouvrage de Masson. Ainsi le terme «avènement» proposé en sous-titre de *L'Image et la parole* n'est-il significativement plus de mise aujourd'hui, où tous les spécialistes s'accordent à considérer que cette «naissance» – plutôt désignée sous le terme de «généralisation» (par exemple chez Martin Barnier)<sup>19</sup> – résulte de pratiques sonores qui remontent aux origines du cinématographe, et même à une période plus ancienne.

17. Alain Masson, *L'Image et la parole. L'Avènement du cinéma parlant*, Paris: La Différence, 1989.

18. Rick Altman, *Silent Film Sound*, New York: Columbia University Press, 2004; Giusy Pisano, *Une archéologie du cinéma sonore*, Paris: CNRS, 2004; Édouard Arnoldy, *Pour une histoire culturelle du cinéma. Au-devant de «scènes filmées», de «films chantants et parlants» et de comédies musicales*, Liège: Éditions du Céfal, 2004. Une somme de ces recherches est proposée dans l'ouvrage collectif dirigé par Giusy Pisano et Valérie Pozner: *Le Muet à la parole. Cinéma et performance à l'aube du XX<sup>e</sup> siècle*, Paris: AFRHC, 2005. Il s'agit des actes d'un colloque dont je rends compte en mettant l'accent sur des phénomènes d'énonciation et d'oralité dans «Le colloque *Le Muet à la parole* de l'auditorium du Louvre», 1895, N° 44, décembre 2004, pp. 95-103.

19. Martin Barnier, *En route vers le parlant. Histoire d'une évolution technologique, économique et esthétique du cinéma (1926-1934)*, Liège: Éditions du Céfal, 2002.

Cette voie a été jusqu'ici peu suivie, puisque force est de constater que les rapports entre l'image et la voix n'ont été étudiés que de manière très épisodique. En France, on pense avant tout à *La Voix au cinéma* de Michel Chion, un ouvrage précurseur qui est cependant dépourvu de perspective historique et offre peu de propositions théoriques, étant essentiellement composé d'une juxtaposition d'analyses non systématiques d'extraits de films dans lesquels la voix est exploitée sur un mode particulier. Son travail « théorique » (qui ne se veut pas détaché de la pratique cinématographique), poursuivi à travers différents ouvrages de synthèse à vocation pédagogique, présente l'avantage de discuter certaines notions ou catégories. La voix n'occupe toutefois pas une place privilégiée dans le travail de Chion, puisqu'il met principalement l'accent sur la matière sonore en général ou sur la musique en particulier. Dans un texte portant significativement sur la « réhabilitation du bruit », Chion notait toutefois que la voix, « thème nouveau qui ne manque pas de fasciner », n'avait donné lieu qu'à quelques essais, contrairement aux textes des dialogues qui ont suscité une pléthore d'analyses<sup>20</sup>. L'absence de modifications dans les rééditions successives de *L'Audio-vision* témoigne d'une certaine stagnation des études dans le domaine.

La voix au cinéma a donc été traitée de façon sommaire, notamment en comparaison avec l'attention traditionnellement portée à la dimension visuelle du cinéma parlant. De plus, les écrits généraux récents sur le son au cinéma ont tendance à exclure ou à minorer la prise en compte de la voix sous prétexte que cette question serait surreprésentée dans les réflexions antérieures. Cette prédominance s'expliquerait par le fait que le dialogue prime sur les autres composantes sonores dans les pratiques cinématographiques courantes. Il faut dire que les théoriciens du cinéma ont tendance à stigmatiser l'étude de la voix comme véhicule d'un texte parce qu'elle témoigne d'une orientation verbocentriste qu'ils associent au courant « structuraliste » des années 1960-1970. Le texte, même *parlé*, devient leur bête noire dans la mesure où il est un « bon objet » d'une certaine théorie du cinéma qui, s'étant édifiée à partir d'autres champs (la sémiotique, la linguistique, l'analyse structurale des récits, etc.), fit du langage verbal un comparant idéal. Lorsque de telles objections à l'étude de la parole ne

20. Michel Chion, *L'Audio-vision*, Paris: Nathan, 1990, p. 124.

reposent pas sur une énième formulation du discours concernant la soi-disant «spécificité cinématographique» de l'image, elles résultent d'une volonté de promouvoir l'examen d'éléments sonores non verbaux. Telle est par exemple la position de Laurent Jullier qui, dans l'état des lieux qu'il dresse en préambule d'un ouvrage sur les sons au cinéma, s'explique comme suit :

«Jusqu'à une date récente, la parade favorite des chercheurs consistait à choisir comme objet d'étude des sons filmiques un peu particuliers, des sons qui font figure de *supports* pour des systèmes pré-existants. En premier lieu la voix, qui sert de support aux mots et permet à l'analyste de convoquer toute une batterie d'outils théoriques élaborés par la linguistique et la narratologie...»<sup>21</sup>

Une grande partie de mon travail relèvera – je le concède et le revendique – d'une telle «parade», car les conséquences des phénomènes vocaux et verbaux sur le dispositif du cinéma et sur le discours filmique ont été à mon sens peu examinées jusqu'ici. En fait, ainsi que le constatait Michel Colin dans un bilan sur l'apport du structuralisme, l'influence de la linguistique s'est plutôt exercée sur les études cinématographiques qui prenaient comme objet des composantes non verbales comme le montage, la prise en compte d'un autre médium permettant de défricher un terrain d'application différent<sup>22</sup>. Somme toute peu théorisées, les relations verbo-iconiques nous donnent l'occasion de rediscuter certains acquis de la narratologie filmique, et plus généralement de l'analyse du film comme discours. En particulier, l'énonciation filmique gagne à être envisagée spécifiquement du point de vue de la voix, et inversement. Il me semble possible d'éviter l'étroitesse qu'y voit Jullier en abordant également cette problématique en dehors de la stricte logique textuelle, par exemple en comparant la voix du cinéma parlant à d'autres spectacles audiovisuels qui font intervenir une «voix vive». En outre, l'intérêt pour les bruits au cinéma manifesté récemment par certains théoriciens de l'audiovisuel, dont Laurent Jullier est à cet égard le principal représentant dans l'espace francophone, est sans conteste suscité par le développement des technologies et des pratiques du

21. Laurent Jullier, *Les Sons au cinéma et à la télévision*, Paris: Armand Colin, 1995, p. 8.

22. Michel Colin, *Langue, film, discours*, Paris: Klincksieck, 1985, p. 27.

*sound design*. Or, cette question touche peu l'histoire du cinéma antérieure aux années 70, époque où sont apparus aux États-Unis, avec la génération de cinéastes comme Coppola ou Lucas, des ingénieurs tels que Walter Murch. Jullier a d'ailleurs consacré précédemment un livre à un type de films dans lesquels les effets sonores priment sur la narration<sup>23</sup>, un cas de figure qui n'invalide pas pour autant, à mon sens, la pertinence d'une analyse narratologique. L'élargissement du champ théorique à toutes les composantes de la bande-son que prône Jullier me paraît certes légitime, mais je ne pense pas que l'on puisse prétendre que le territoire couvert par la question de la voix au cinéma a été totalement exploré, et cela même dans les films qui représentent pour lui le parangon d'une nouvelle exploitation du caractère spectaculaire des sons non soumis au langage articulé, à l'exemple de la série des *Star Wars*<sup>24</sup>. D'ailleurs, l'examen de la voix conduit parfois à analyser des éléments sonores comme la musique ou les bruits qui apparaissent concurremment à la voix, ou qui ont été pensés en des termes qu'il est intéressant de confronter à la fonction de cette dernière. Par exemple, la voix est soumise, à l'instar d'autres sons filmiques, aux techniques d'enregistrement et de reproduction qui la distinguent des voix «réelles» perçues dans notre quotidien. Les implications de la technique phonographique qui seront discutées concernent donc également, à de nombreux égards, l'univers sonore dans son ensemble. De plus, la parole interagit souvent intimement avec des occurrences non verbales comme la musique, notamment dans le cas «mixte» du chant dont il sera question sous l'angle de son intégration au récit filmique.

## ÉNONCIATION ET VOIX-OVER

La voix-over constitue en fait le point de mire de mon étude puisque la conception graduelle du désancrage de la voix par rapport à l'image culmine dans ce procédé. En la comparant à

23. Laurent Jullier, *L'Écran post-moderne*, Paris: L'Harmattan, 1997.

24. En effet, dans les deux trilogies de *La Guerre des étoiles* (1977-2004), le véhicule principal du récit est indubitablement le verbal, puisque la structure de ces films procède, sur le modèle des films de guerre classiques, d'une alternance de scènes dialoguées faisant intervenir différentes «forces» en présence. En mettant exclusivement l'accent sur le travail effectué sur les bruits, on occulte à mon avis l'importance que ces films accordent aux voix proprement dites. En effet, les langues imaginaires, certaines oppositions au niveau de la vocalité (voix humaines *vs.* mécaniques *vs.* animales, accent anglais *vs.* *slang* américain) ou la récurrence de voix transmises par des moyens futuristes, émises par des visages dissimulés derrière des masques figés ou même intériorisées grâce à des phénomènes «paranormaux», font de la voix et de la parole des éléments-clés des *Star Wars*.

d'autres phénomènes vocaux, j'essaierai de saisir sa spécificité du point de vue de la production du sens, de l'élaboration diégétique et du réglage pragmatique qu'elle instaure. Il faut noter que deux ouvrages existent déjà sur la question, l'un de Sarah Kozloff datant de la fin des années 80<sup>25</sup>, l'autre de Jean Châteaupert, paru dans la seconde moitié des années 90<sup>26</sup>. De la restriction de leur objet au seul procédé de la voix-over découle cependant une approche passablement différente de la mienne, puisque je projette d'étudier divers types de manifestations vocales en remontant également en deçà des débuts « officiels » du parlant. Châteaupert aborde en effet bien moins souvent la question du bonimenteur que la préface de Roger Odin ne pourrait le laisser présager<sup>27</sup>, car sa réflexion ne tient pas compte de l'opposition entre voix vive et voix enregistrée; quant à Kozloff, qui fait porter son étude sur un objet plus ciblé (le cinéma américain de fiction), elle manifeste une attention certaine envers d'autres formes vocales mais, dans son chapitre consacré aux « ancêtres », le bonimenteur occupe à peine plus d'une page (ce qui s'explique par l'état encore embryonnaire de la recherche sur le boniment à l'époque de la rédaction de l'ouvrage). Globalement, ma perspective est plus proche de celle de Châteaupert, notamment en ce qui concerne son optique pragmatique (bien qu'il ne sorte jamais de l'institution du cinéma ni ne prenne en compte le « dispositif ») et sa conception de la diégèse filmique, inspirée de la théorie des mondes possibles issue du champ de l'analyse littéraire. Par contre, je m'éloignerai considérablement du modèle énonciatif proposé dans son ouvrage, non pas pour le nier en bloc, mais pour l'intégrer à des bases différentes résultant d'une prise en compte de la spécificité de la voix enregistrée. Ces questions d'énonciation filmique seront discutées en détail au cours du chapitre VI, lieu de convergence de la plupart des conclusions théoriques de ce travail, où je repartirai notamment des premiers écrits d'Albert Laffay, pour qui la voix-over constituait un objet privilégié d'une approche visant à s'interroger sur la place de ce qu'il appelait le « grand imagier ». On verra d'ailleurs que la référence au cas particulier de la voix-over est récurrente chez les théoriciens du cinéma qui se sont penchés sur la question de la « voix

25. Sarah Kozloff, *Invisible Storytellers. Voice-over Narration in American Fiction Film*, Berkeley; Los Angeles; London: University of California Press, 1988.

26. J. Châteaupert, *Des mots à l'image...*, op. cit.

27. *Ibid.*, p. 13.

narrative», notion provenant du champ littéraire qui mérite à mon sens d'être rediscutée comparativement au cinéma, où la «voix» acquiert une existence concrète.

En inscrivant la problématique de l'énonciation filmique dans la dialectique de l'humain et de la machine qui traverse l'ensemble de mon travail, je tenterai de nuancer le postulat anthropomorphique de Kozloff et de Châteauvert qui, selon moi, trop étroitement redevable des théories littéraires, conduit à occulter tout le pan machinique de la production de la représentation sonore. Pour ce faire, je m'inspirerai des principes du dernier opus de Christian Metz significativement intitulé *L'Énonciation impersonnelle ou le site du film*, dont l'auteur n'a à mon sens pas pleinement exploité le potentiel théorique. C'est justement parce que toute voix est nécessairement associée à une «personne» que le caractère «impersonnel» de l'énonciation prend une importance considérable. L'humain intrinsèquement évincé du dispositif du cinéma parlant fait retour dans le film à travers sa mise en scène vocale.

### **LE BONIMENTEUR DU CINEMA DES PREMIERS TEMPS, UNE PRATIQUE ORALE**

La voix présente la particularité de n'exiger fondamentalement aucune médiation technologique pour intervenir au «cinéma». Elle y a donc toujours été présente – ne serait-ce qu'à l'extérieur de la baraque foraine pour attirer les spectateurs – mais elle fut également le lieu d'importantes mutations technologiques, principalement liées à l'enregistrement et à la diffusion phonographiques, ainsi qu'aux divers procédés de synchronisation avec l'image. Afin d'étudier la voix dans un réseau plus large de pratiques audiovisuelles, je mettrai l'accent sur les spectacles vivants (notamment le vaudeville américain), lieu d'exercice d'une «voix-attraction» que l'on opposera au régime de la voix-narration régissant la plupart des voix-over. On peut en outre distinguer deux types de manifestation vocale en fonction de leurs conditions matérielles de profération : la voix *vive* et la voix fixée. C'est pourquoi je propose de développer l'idée émise par François Albera selon laquelle il y aurait une «appartenance native du cinéma au domaine de l'oralité au sens large, celui qu'a pu définir Paul Zumthor»<sup>28</sup>, appartenance par la

28. François Albera, «Ouverture», 1895, N° 38 («Musique!»), 2002.



suite infléchie dans le sens contraire de l'«inscription» phonographique, en envisageant la question de la parole, profondément traversée par l'opposition de l'oral et de l'écrit, dans le cadre des travaux menés par Paul Zumthor sur la «performance». Dans le champ du cinéma, Germain Lacasse s'est déjà référé aux écrits du médiéviste pour discuter certaines particularités de la pratique du bonimenteur<sup>29</sup>, mais il m'a semblé fructueux d'y recourir de façon plus approfondie en examinant les conséquences de l'intervention de la parole orale dans le «dispositif cinématographique». Par ailleurs, la comparaison menée ici entre l'oralité et les conditions techniques du «parlant» exige de nuancer ou d'adapter les conceptions de Zumthor.

L'accompagnement verbal des projections cinématographiques du début du XX<sup>e</sup> siècle est désormais un fait avéré qui a donné lieu, principalement ces dix dernières années, à plusieurs études dont les principaux résultats ont été rassemblés en 1996 dans le numéro 22 de la revue *Iris* dirigé par Gaudreault et Lacasse, et dans l'ouvrage de ce dernier, *Le Bonimenteur des vues animées*. Grâce à une comparaison entre cinéma parlant et «cinéma» parlé, à l'emprunt de certaines théories linguistiques et à la prise en compte de la place allouée à la voix au sein d'un dispositif, je pense pouvoir offrir un éclairage nouveau sur le rôle de l'accompagnateur de vues et sur les conclusions plus générales que l'on peut tirer de son activité.

Pour qualifier ce locuteur dont la voix se faisait entendre durant les projections des «premiers temps», j'opterai pour le terme «bonimenteur» qui, indépendamment de l'usage spécifique que l'on pouvait en faire à l'époque ou que lui confèrent certains historiens (comparativement au «conférencier», au «bonisseur»<sup>30</sup>, etc.), fera pour moi office d'hyperonyme, son utilisation s'étant désormais généralisée dans la plupart des écrits francophones sur la question. André Gaudreault est récemment revenu sur cette problématique dans le cadre d'une recherche menée avec François Albera pour démontrer combien l'instabilité terminologique avait mené à l'évacuation de la figure du «bonimenteur» au sein de

29. Germain Lacasse, *Le Bonimenteur de vues animées. Le cinéma «muet» entre tradition et modernité*, Québec; Paris: Nota Bene, Méridiens Klincksieck, 2000.

30. En raison des variations notables que l'on constate dans les textes d'époque, il me paraît plus clair de réserver le terme «bonisseur» au collaborateur du spectacle cinématographique qui se trouve à l'extérieur de la baraque et harangue la foule. Il est donc distinct du «bonimenteur» qui, présent dans la salle, accompagne la projection.

l'historiographie française du cinéma<sup>31</sup>. Il n'est en effet pas toujours probant de se référer uniquement aux termes utilisés à l'époque, car le mot «bonimenteur» pouvait être écarté, selon l'institution concernée, en raison des connotations liées à la nature populaire et à la fonction strictement divertissante du spectacle. Arlaud donne un exemple de telles déterminations lorsqu'il évoque un type particulier d'exploitants, les «ambulants»: «Il ne faut jamais, devant les ambulants ou leur descendance, parler de bonimenteur. Des bonimenteurs, il y en avait à Paris, chez Dufayel, les forains font le boniment, mais les Régnault (faisaient la conférence.»<sup>32</sup> Il semble donc bien que, dans le discours même des «bonimenteurs», cette question terminologique ait participé d'une différenciation entre ambulants et forains qui, pour ceux-là, constituait un élément de valorisation culturelle de leur pratique. L'analyse des histoires francophones du cinéma conduit Albera et Gaudreault à affirmer qu'il vaut mieux trancher pour le terme «bonimenteur» plutôt que de sous-estimer le phénomène. C'est ce que je ferai ici en proposant en outre une adjectivation du terme («bonimentoriel») utile pour qualifier notamment les processus de *médiation* qu'implique selon moi l'insertion du bonimenteur dans le dispositif.

Les études portant sur le bonimenteur se sont majoritairement orientées, parce qu'il était nécessaire d'opérer un premier défrichage (qui reste encore largement à faire) dans un terrain presque totalement vierge, vers des considérations qui ont trait au caractère local ou national du phénomène. Ainsi le *benshi* s'inscrit-il dans une tradition culturelle proprement nipponne, et le bonimenteur québécois qui occupe prioritairement Lacasse se fait le porte-parole d'une minorité. J'aimerais quant à moi sortir du cadre d'un examen (voire d'une défense) des particularismes pour concevoir les implications générales de la pratique du boniment au sein d'une culture de masse en pleine constitution, celle de l'audiovisuel. Les hypothèses de Germain Lacasse ne manquent pas d'intérêt, mais elles méritent à mon sens d'être envisagées dans un spectre plus vaste de possibilités. Paradoxalement, Lacasse fait une

31. François Albera et André Gaudreault, «Apparition, disparition et escamotage du bonimenteur dans l'historiographie française du cinéma», in G. Pisano et V. Pozner (dir.), *Le Muet à la parole...*, op. cit., pp. 167-198.

32. Rodolphe-Maurice Arlaud, *Cinéma-Bouffe. Le Cinéma et ses gens*, Paris: Éditions Jacques Melot, 1945, p. 76.

généralité de la pratique de résistance, alors qu'il la définit justement par opposition au mode dominant de consommation des images, comme si ce dernier excluait nécessairement le boniment. Mieux vaut à mon avis considérer les rapports « conflictuels » entre la parole vive et le « film », majoritaires dans certaines régions « coloniales », comme l'une des postures pour lesquelles le bonimenteur peut opter, plutôt que comme une période spécifique de l'histoire de cette pratique. Il importe à mon sens de ne pas réduire la diversité des interactions entre l'image, la voix et le public à une fonction sociopolitique définie, sous peine de ne pouvoir rendre compte des différentes modalités qui ont régi l'usage de la voix lors de projections bonimentées. Il est néanmoins pertinent d'affirmer à un niveau plus général que toute pratique instaurée au niveau de l'exploitation constitue une source potentielle de déviation du sens et des valeurs véhiculées par les films. Mary Carbine a par exemple montré combien la musique jazz jouée dans les salles du « Black Metropolis » de Chicago permettait à une minorité d'investir le texte filmique d'une dynamique totalement différente de celle originellement prévue, offrant une lecture à rebrousse-voix qui pouvait passer par l'ignorance de la dimension narrative<sup>33</sup>. La musique d'accompagnement s'oppose à la machinerie au même titre que le boniment. Il s'agit néanmoins de considérer les implications plus spécifiques de la vocalité et du verbal liées à l'activité bonimentorielle. À cette fin, je me référerai à la réflexion menée par Paul Zumthor à propos de l'« oralité » tout en la dégagant de ses pré-supposés critiques qui ont passablement influencé le travail de Germain Lacasse (ayant vécu en Suisse romande et au Québec, Zumthor est comme lui un habitué des minorités). En effet, il importe au médiéviste d'opter pour un objet (la poésie orale) qui lui permet d'aborder une époque antérieure (ou des pratiques extérieures) à ce qu'il appelle « la culture d'origine européenne liée à la civilisation technologique et en voie d'universalisation rapide et brutale »<sup>34</sup>. Lorsque Zumthor évoque incidemment la présence de l'oralité dans les *mass media*, ce n'est significativement pas sans suggérer une certaine dénaturation occasionnée par les modifications (qu'il discute fort peu) induites par la phonographie (il qualifie par exemple la technologie du son de « quincaillerie

33. Mary Carbine, « The Finest Outside the Loop »: Motion Picture Exhibition in Chicago's Black Metropolis, 1905-1928 », *Camera obscura*, N° 23, 1990, p. 31.

34. Paul Zumthor, *Introduction à la poésie orale*, Paris: Seuil, 1985, p. 65.

industrielle)»<sup>35</sup>. Néanmoins, ce défenseur de l'oralité précise ailleurs, ainsi que l'avait fait Walter Ong<sup>36</sup>, que la voix a trouvé selon lui au XX<sup>e</sup> siècle un nouveau moyen d'affirmer son autorité en affichant une véritable continuité avec les pratiques primitives de toutes les cultures humaines<sup>37</sup>. En dépit du préjugé «anti-machinique» de son auteur, cette poétique de l'oralité se prête à une application au «cinéma» parlé. Elle permet en effet de concevoir sur un plan très général des outils descriptifs indépendamment du contexte idéologique et de l'objet initial de la recherche<sup>38</sup>. Ses réflexions nourriront par conséquent les thèses avancées ici à propos du pôle «humain» des manifestations vocales au cinéma. Germain Lacasse rappelle que le cinéma relève pour Zumthor pleinement de la catégorie générique de la «Poésie orale» (puisque le texte n'y existe qu'en emploi, sous la forme d'une voix), et note que cette appartenance vaut d'autant plus pour le cinéma des premiers temps<sup>39</sup>. La théorie exposée par Zumthor permet plus largement à la fois de tisser des liens entre boniment et voix du parlant (autour du concept de «vocalité») et d'opérer des distinctions (autour de la notion opératoire de «performance»). C'est pourquoi je ferai dialoguer certaines de ses réflexions avec les thèses des historiens du cinéma qui se sont penchés sur la question du bonimenteur.

Zumthor distingue trois types d'oralité: l'oralité *primaire*, qui ne peut concerner un médium apparu dans une société où écriture et oralité coexistaient déjà; l'oralité *mixte*, qui pourrait s'appliquer aux premières années du cinématographe (puisque l'influence de l'écrit sur la performance orale était moindre), si le cinéma n'était lui-même le produit d'une culture «lettrée» profondément marquée par l'expression écrite; enfin, l'oralité *seconde* qui, recomposée à partir de l'écriture, caractérise le cinéma, et d'autant plus le cinéma parlant<sup>40</sup>. On pourrait ajouter que le degré de cette secondarisation de l'oralité connaît des variations en fonction de l'existence ou non d'une source écrite, et du rapport que le locuteur

35. *Ibid.*, p. 28.

36. Walter J. Ong, *The Presence of the Word*, New Haven; London: Yale University Press, 1967, p. 88.

37. Paul Zumthor, «Le geste et la voix», *Hors Cadre*, N° 3 («Voix off»), 1985, p. 73.

38. Par la suite, Zumthor a d'ailleurs lui-même souligné la nature interdisciplinaire de sa démarche (voir *Performance, réception, lecture*, Longueuil: Les Éditions du Préambule, 1990, pp. 9-11).

39. G. Lacasse, *Le Bonimenteur...*, *op. cit.*, p. 129.

40. Paul Zumthor, *La Lettre et la voix*, Paris: Seuil, 1987, pp. 18-19.

entretient avec cette dernière. Comme l'explique Le Men, ces possibilités s'offraient déjà au montreur de lanterne magique qui pouvait « lire à haute voix le texte des plaques ou celui du livret de colportage, ou encore inventer des variations de conteur sur le canevas offert par les plaques »<sup>41</sup>. Le bonimenteur de cinéma, auquel les maisons de production mettaient parfois à disposition des textes d'accompagnement, était confronté aux mêmes choix.

Plus les contraintes reposant sur la performance orale sont importantes, plus la représentation audiovisuelle s'apparente à une communication qui s'opère sur le mode de *l'écrit*. Saisir le degré d'appartenance de certaines pratiques (pré)cinématographiques au domaine de l'oralité, tel sera l'un des objectifs de cette étude. Je pense en effet qu'en réfléchissant dans un premier temps à la question de l'oralité, on est ensuite plus armé sur le plan théorique pour aborder le statut de la parole enregistrée du cinéma parlant. Le rapprochement entre boniment et voix-over doit néanmoins être nuancé en tenant compte des spécificités de la « médiation bonimentorielle » qu'il s'agira de définir à partir de théories linguistiques et communicationnelles. Lorsque André Gaudreault qualifiait le bonimenteur de « commentateur en *voix over*, mi-visualisé, mi-acousmatique »<sup>42</sup>, il empruntait des catégories développées dans le cadre du cinéma parlant pour tracer une filiation et ouvrir la voie à une réflexion sur la prestation du bonimenteur. À partir du cinéma bonimenté, Gaudreault propose d'essayer « de comprendre la place que viendra occuper, ultérieurement, la « voix » narrative du narrateur en « voix » *off*, auquel allait donner son essor le cinéma parlant »<sup>43</sup>. C'est dans cette perspective, et avec le même dessein sous-jacent d'examiner les implications de l'insertion d'un énonciateur humain dans un dispositif « machinique », que je discuterai la question du « cinéma » bonimenté.

Du boniment à la voix-over, il y a cependant plus qu'un simple pas : dans un mouvement général qui part du *contexte* historique et culturel (domaine de l'exploitation cinématographique qui accueille le boniment) pour aboutir au *texte* filmique (la voix-over comme composante discursive), je tenterai de saisir les divers paramètres

41. Ségolène Le Men, *Lanterne magique, tableau transparent*, Paris: Réunion des Musées nationaux, 1995, p. 38.

42. André Gaudreault, *Du littéraire au filmique*, Québec; Paris: Nota Bene, Armand Colin, 1999 [1<sup>re</sup> édition 1988], p. 147.

43. *Ibid.*, p. 143.

régissant le fonctionnement respectif de ces deux types de manifestations vocales caractérisés par une forme (différente) d'extériorité à la diégèse visualisée. L'application de la notion d'« attraction » aux régimes vocaux permet en outre de prolonger certaines propositions développées dans les années 80 sous l'égide de la nouvelle approche instaurée grâce au colloque de Brighton (1978), qui permit non seulement de réhabiliter le cinéma des premiers temps en affirmant sa singularité esthétique (invalidant ainsi la vision téléologique d'un « primitivisme » qui serait dépourvu d'intérêt propre), mais qui consacra également, avec des auteurs comme Gunning ou Gaudreault, une rencontre féconde de l'*historique* et du *théorique* que l'on observe aujourd'hui encore trop rarement dans le champ passablement compartimenté des études académiques (françaises) sur le cinéma. En fondant la plupart de mes propositions théoriques sur des constats historiques, je m'emploierai à décloisonner autant les objets – la voix traversant quantité de pratiques spectaculaires – que les approches (historiographie, sémiologie, narratologie, sémantique, histoire des techniques, etc.), espérant ainsi défricher, au gré des croisements et de nécessaires ajustements, quelques zones nouvelles du domaine de recherche qu'ouvre la riche problématique des voix au cinéma.

# I. VOIX VIVE ET DISPOSITIFS (PRE)CINEMATOGRAFHIQUES

## I.1 PREALABLES THEORIQUES: VOIX ET LOCUTEUR DANS LES DISPOSITIFS AUDIOVISUELS

Si l'on prend en compte, comme tout un pan de l'archéologie du cinéma s'applique à le faire depuis quelque temps, différents cas d'utilisation du «son avant le parlant», on constate que la voix interagissait déjà avec les autres matières de l'expression sonore (bruits et musique). En focalisant l'attention sur la seule composante vocale, on peut penser autrement cette période, et avancer certains éléments de réflexion sur la voix dans le cinéma parlé et parlant. C'est pourquoi je discuterai les implications de dispositifs pré-cinématographiques ou de contextes spectaculaires ultérieurs qui recourent à la coprésence de la voix et de l'image (chapitres I et V). Parallèlement à ces incursions historiques, j'établirai un cadre théorique général autour de la notion de «dispositif», qui s'avérera notamment utile pour commenter les diverses fonctions du bonimenteur à l'époque du cinéma des premiers temps et, de façon plus large, les liens qui s'instaurent entre un spectacle et son public. L'examen des interactions entre l'image et la parole gagne à être ancré dans un contexte plus large, celui de l'*audiovisuel*. Dans une perspective historique, cette notion peut être appliquée à une période bien plus étendue que celle à laquelle elle est historiquement attachée (d'après le *Petit Robert*, le terme aurait fait son apparition en 1947). On peut en effet remonter en amont de la phase de généralisation des moyens de communication de masse qui s'est opérée au milieu du XX<sup>e</sup> siècle. C'est l'avis de l'historien Jacques Perriault, dont les recherches sur l'usage de la lanterne magique et du phonographe prennent appui sur l'audiovisuel contemporain<sup>44</sup>.

44. Jacques Perriault, *Mémoires de l'ombre et du son. Une archéologie de l'audio-visuel*, Paris: Flammarion, 1981.

Il ne faut pas voir là l'indice de cette téléologie décriée haut et fort par les Nouveaux historiens, car l'analyse de techniques anciennes offre au contraire un cadre théorique permettant de penser d'autres phénomènes plus récents, les « nouvelles » technologies jouant à certains égards des dispositifs antérieurs. Cette reconduction de principes plus anciens vaut d'autant plus pour la voix, facteur humain dont le traitement présente une certaine constance en dépit de l'évolution technologique.

Le projet de catégorisation et de conceptualisation des dispositifs visuels développé par Maria Tortajada et François Albera me semble fournir une base de réflexion intéressante pour aborder plus largement l'audiovisuel<sup>45</sup>. Ces auteurs proposent de fonder l'analyse d'un dispositif sur l'articulation de trois pôles en perpétuelle interaction: le *spectateur*, la *machinerie* et la *représentation*. Même si certaines occurrences sonores peuvent sembler dépourvues de toute médiation assimilable à une « machinerie » – comme la voix d'un commentateur de vues, qui ne se distingue pas dans sa nature des voix entendues dans le quotidien – ce modèle incite à étudier les fonctions spécifiques qu'elles endossent lorsqu'elles sont intégrées à un dispositif spectaculaire. Dès lors qu'il y a coprésence de la production d'une image et de l'émission d'une voix, cette dernière doit être considérée comme partie intégrante du dispositif. Qu'elle soit reproduite et amplifiée de façon mécanique ou électrique, ou prononcée « naturellement », la voix contribue à l'assignation d'un certain positionnement spectatoriel. La notion de « dispositif » peut donc à mon sens être élargie au-delà (ou en deçà) de la sphère technologique à la coexistence, dans un espace commun possédant des caractéristiques données, d'un spectateur et d'une instance de production. Cette dernière peut englober divers éléments auxquels je me référerai en parlant d'*espace de la production de la représentation audiovisuelle*. Il comprend autant des instances humaines (musiciens, chanteurs, bonimenteur, bruiteurs, chef d'orchestre, projectionniste, etc.) que des machines (appareil de projection, machine à bruit, phonographe, mécanisme de synchronisation, etc.). À l'intérieur de cet espace s'établissent de nombreuses relations verticales dont l'objectif est généralement de

45. François Albera et Maria Tortajada, « L'Épistémè '1900 », in André Gaudreault, Catherine Russell et Pierre Véronneau (dir.), *Le Cinématographe, nouvelle technologie du XX<sup>e</sup> siècle*, Lausanne: Payot, 2004.



parvenir à une «synchronisation» au sens large qui, selon les époques ou les pratiques, a pu conduire à engager des moyens humains, mécaniques ou semi-mécaniques. La description suivante de Rodolphe-Maurice Arlaud, qui raconte ses souvenirs de projections bonimentées, est significative de l'importance des interactions entre les différentes instances de l'espace de la production :

«Devant le conférencier, il y avait maintenant une rangée de petits boutons et c'était, entre la cabine et lui, un mystérieux et silencieux dialogue ponctué de lumières rouge, blanche, verte ou bleue. Cela voulait dire: «Passe une vue fixe, passe le film.» Ou bien cela devait arrêter l'orchestre, le faire repartir, changer de partition.»<sup>46</sup>

Selon ce témoignage qui rend compte d'une configuration et d'un fonctionnement spécifique de l'espace de production, le bonimenteur se trouve aux commandes des relations verticales internes à la machinerie. On voit que les pratiques hétérogènes des «sons du muet», plus ou moins artisanales et novatrices, nécessitent d'élargir le pôle de la machinerie à la présence de l'humain, dont le principal représentant est une figure visible et audible de tous: le bonimenteur. C'est pourquoi cette instance joue le rôle d'une sorte de passeur qui, pour l'audiospectateur, définit un accès donné à la représentation visuelle d'une part, à la «machinerie» d'autre part. Elle ne correspond à aucun des trois pôles exactement, mais intervient entre chacun de ceux-ci. Cette position intermédiaire résulte des caractéristiques suivantes: bien qu'humain, le bonimenteur est une source de production comparable à la machinerie, et ce d'autant plus qu'il peut être remplacé par un phonographe, ou intervenir conjointement avec ce dernier; il entretient une relation directe avec ceux qui, au même titre que lui, regardent les images; il fait intervenir son corps pour donner plus de poids à ses dires, créant ainsi une représentation visuelle «concurrente» de celle proposée par l'image projetée.

La situation où un locuteur s'adresse à des auditeurs implique certaines variables (en termes de proxémique, d'interactivité, etc.) dont la combinaison mutuelle définit un dispositif donné, fût-il «minimal». Par exemple, les «récits au coin du feu» impliquent

46. R.-M. Arlaud, *op. cit.*, p. 69.

une disposition centrée autour d'un élément qui stimule la participation d'autres sens que l'ouïe (la vue par la lumière, le toucher indirectement par la chaleur), tout en reléguant l'environnement dans l'obscurité. En considérant des « situations d'énonciation » de ce type non pas seulement au niveau du contenu du texte proféré, mais comme des dispositifs, on peut tenter de formuler quelques propositions théoriques concernant les diverses manifestations de l'oralité. Ces postulats ne sont pas sans incidences sur la manière dont j'envisagerai la place de la voix dans le cinéma des premiers temps: au lieu d'affirmer, comme Tom Gunning, que « le conférencier révèle une fissure à l'intérieur du cinéma comme dispositif »<sup>47</sup>, je préfère appréhender cette pratique comme un élément solidaire de la projection des images. En effet, en contribuant à configurer les relations entre les trois pôles machinerie/représentation/spectateur, le boniment donne lieu à un dispositif spécifique que l'on appellera « *cinéma* » parlé. Certes, bien que se situant du côté de la machinerie en tant que facteur de production de la représentation, le bonimenteur n'est pas assimilable à un élément « mécanique », puisque son rôle consiste précisément à contrebalancer, par une plus-value d'humanité, la technicité inhérente à la composante visuelle du dispositif. Dans notre système tripolaire, cette instance humaine occupe en fait une place variable, mais procède dans tous les cas à une mise en relation des différents composants du dispositif. Selon le type de médiation dont elle procède, la voix est happée plus ou moins fortement dans l'orbite de l'un de ses trois pôles.

Le modèle tripartite machinerie (et/ou instance humaine)/spectateur/représentation valable pour la production d'images s'applique également aux dispositifs sonores. Du Kinétophone (écoute individuelle avec paiement pour une durée limitée) à la retransmission radiophonique différée, en passant par une écoute à distance en simultané (possible dès l'invention du Théâtophone en 1881, qui proposait une écoute individuelle), par les divers types de phonographes à usage domestique ou présentés par une instance humaine, chaque système induit un dispositif spécifique. Dans cette étude, je me restreindrai aux dispositifs qui convoquent concurremment l'image et la voix, et j'écarterai tous les spectacles proprement scéniques reposant sur une représentation en actes

47. Tom Gunning, « The Scene of Speaking: Two Decades of Discovering the Film Lecturer », *Iris*, N° 27, 1999, p. 78.

effectuée par des acteurs visibles, car ces cas de figure exigeraient de prendre en compte un spectre de pratiques trop large. Par ailleurs, je m'en tiendrai aux dispositifs qui présentent, à l'un au moins des deux versants de la représentation (auditif ou visuel), une composante mécanique, celle-ci caractérisant toute projection cinématographique. Ces objets permettront de discuter les diverses modalités d'intégration de la voix, attribut de l'humain, dans un dispositif partiellement ou majoritairement basé sur des facteurs techniques ou technologiques.

La corrélation entre la (re)production mécanique de sons et la monstration d'images exige une conception bipartite de la machinerie et de la représentation (voire du destinataire qui, confronté à une dissociation sensorielle, devient un audio/spectateur). S'instaure alors ce que je propose d'appeler une *relation verticale* entre les composantes visuelles et sonores à l'intérieur de la machinerie ou de la représentation, et une *relation horizontale* entre ce qui appartient à la machinerie et ce qui relève de la représentation. Le premier type de relations concerne le couplage des appareils d'émission sonore avec la projection visuelle et, au niveau de la représentation, les liens qui s'établissent *de facto* en cas de cooccurrence d'images et de sons. La relation horizontale, généralement moins perceptible au cinéma du fait qu'elle constitue un obstacle à l'illusion référentielle, se manifeste par exemple lorsque l'installation sonore influence ostensiblement notre perception de l'espace diégétique. Il faut aussi considérer la nature représentationnelle des sons eux-mêmes : comme l'explique Rick Altman, les sons enregistrés ont pour fonction de «représenter les événements sonores plutôt que de les reproduire»<sup>48</sup>. Cette prise en compte de la nature représentationnelle des sons enregistrés innovera l'ensemble de ce travail, redevable à cet égard des réflexions menées principalement par Alan Williams puis James Lastra. On postulera donc que les occurrences sonores produites par des moyens mécaniques relèvent donc déjà, en elles-mêmes, du pôle de la représentation.

Il faut souligner que le «perçu» joue un rôle primordial dans les phénomènes qui impliquent la représentation, puisque son élaboration transite nécessairement par le spectateur et/ou l'auditeur. Le processus d'institutionnalisation du «cinéma» provoqua, en privilégiant

48. Rick Altman, «The Material Heterogeneity of Recorded Sound», in *Sound Theory, Sound Practice*, op. cit., p. 29.

la narration et un univers diégétique clos, une «verticalisation» des relations horizontales: la pratique courante du synchronisme s'inscrit dans cette tendance puisqu'elle vise à gommer la relation horizontale en la déplaçant illusoirement sur le seul plan de la représentation. À des degrés variables, ces deux types de relation peuvent être soit occultés, soit exhibés par le dispositif.

On peut se demander ce qu'il advient de ces relations lors d'une intervention humaine. Étant donné que le locuteur introduit également une composante visuelle seconde, elles se voient nécessairement complexifiées. En effet, un présentateur non dissimulé derrière l'écran ou dans l'obscurité se donne à voir, regarde le spectateur et peut, par son jeu, sa gestuelle, sa mimique ou son habillement, créer un univers référentiel propre qui est susceptible de prolonger, de contredire, de ridiculiser, etc. celui du film. Dans son exposé sur l'art oratoire du conférencier qui accompagnait les projections de lanterne magique, G.-M. Coissac précisait: «La parole n'est pas tout, en effet; l'expression de la physionomie, le geste qui l'accompagne lui donnent plus d'énergie et de sens.»<sup>49</sup>

Dans le cadre de l'institution cinématographique où l'image (c'est-à-dire le «film» comme unité) constitue l'élément central du spectacle, le rôle du bonimenteur est cependant inféodé à ce qui est montré (avec toutefois la latitude propre au «cinéma *parlé*»), car il est engagé pour commenter les films, non pour s'adresser librement à l'auditoire comme le faisait le bonisseur à l'entrée de la salle. Les paroles du bonimenteur s'inscrivent donc volontairement dans un rapport de complémentarité à la «diégèse» filmique. À ce niveau, le fonctionnement signifiant qui sous-tend la représentation est identique au phénomène de coréférence produit par le procédé de la voix-over, même si la synchronisation n'est pas aussi précise. Par contre, la présence physique du bonimenteur, instance désolidarisée du lieu de la production mécanique des images, peut donner lieu à une exhibition de ce dernier. On observe ce phénomène avec cet ancêtre du bonimenteur de cinéma que fut le conférencier des spectacles de lanterne magique. Par exemple, certaines recommandations publiées par la compagnie new-yorkaise de production et de diffusion de plaques York and Son avaient trait à l'insertion du locuteur dans le dispositif. Dans ce manuel, on

49. G.-M. Coissac, *Manuel pratique du conférencier-projectionniste*, Paris: La Bonne Presse, 1908, p. 179.

conseille aux conférenciers amateurs «de ne pas adopter la façon trop commune de faire signe à l'opérateur en frappant avec la baguette, et simultanément en donnant des indications avec la voix»<sup>50</sup>. Il s'agit donc de supprimer la «redondance» des signaux et d'écartier toute référence à la relation verticale entre le conférencier et le projectionniste, celle-ci influant, d'après les auteurs, sur la relation horizontale au spectateur. Coissac prend à cet égard un exemple concret afin de valoriser le système de projection par transparence, qui permet à l'opérateur de prendre place derrière l'écran. L'inconvénient de la projection traditionnelle avec projecteur situé derrière (ou parmi) le public réside selon lui dans le fait que «l'opérateur est généralement entouré d'un grand nombre de spectateurs, dont la curiosité est éveillée par le plus petit détail des manipulations; ils enlèvent ainsi de l'attention que réclame le professeur ou le conférencier»<sup>51</sup>. Cet élément est considéré comme perturbateur – du point de vue de l'information transmise verbalement, non de l'immersion du spectateur dans le monde de l'image, qui est secondaire pour Coissac – parce qu'il résulte d'une exhibition de la relation horizontale entre la représentation et la machinerie. «L'espace de la production» s'affiche également par l'usage de la baguette, ce prolongement physique (et donc, à un niveau minimal, «technique») du locuteur dans l'espace physique de l'écran. Si, en raison de la soumission de l'image au verbal qu'il semble impliquer, on associe généralement cet instrument au contexte de la projection pédagogique – Coissac conseille par exemple au «conférencier» d'en faire usage<sup>52</sup> – il n'est pas exclu que le bonimenteur de cinéma l'ait également utilisée, ne serait-ce que pour se donner des airs de «conférencier». À l'instar de Jean Keim, qui fait mention de «la longue baguette» du bonimenteur<sup>53</sup>, R.-M. Arlaud se souvient d'un bonimenteur engagé par le directeur d'un établissement forain dont il décrit l'activité en ces termes :

«Celui-ci venait tout droit du vieux montreur d'images. Il en avait même la baguette. Il commentait ce que tout le monde voyait, frappait de grands coups sur la toile, s'échauffait plus que

50. «Preliminary Hints to Amateur Lecturers» (vers 1880), cité par R. Crangle, *art. cit.*, p. 45.

51. G.-M. Coissac, *op. cit.*, pp. 198-199.

52. *Ibid.*, pp. 182-183.

53. Jean A. Keim, *Un nouvel art. Le Cinéma sonore*, Paris: Albin Michel, 1947, p. 41.

les acteurs. [...] La baguette désignait Paul. «Elle l'aperçoit.» La baguette soulignait la terreur de Juliette. «Elle se retourne!»<sup>54</sup>

Cet accessoire permet d'expliciter et de visualiser les phénomènes de coréférence de la parole et de l'image, tout en soulignant la présence d'une instance humaine. Sa fonction explicative ne porte pas uniquement sur le référent visuel, mais sur la lecture qu'il faut en faire, et de la sorte oriente le regard du spectateur. André Gaudreault a d'ailleurs rapproché le bonimenteur de l'«admoniteur» de certaines œuvres picturales qui, le plus souvent, tend le doigt pour indiquer au spectateur l'élément central de la composition<sup>55</sup>. Dans ce cas, la fonction narrative peut être intimement liée à l'explication du contenu de la scène montrée, ce qui nous incite à décroquer certains types d'organisation discursive. Jean Châteauvert a relevé que certains personnages qui apparaissent dans les films du début du siècle offraient, grâce à leur mimique et à leur gestuelle, une sorte de discours adressé au spectateur, «discours qui contribue à éclaircir certaines dimensions du récit»<sup>56</sup>. Contrairement à ceux-ci, le conférencier à baguette appartient néanmoins ostensiblement au même espace-temps que son public, ainsi que le rappelle la tridimensionnalité de son instrument qui se détache sur le fond des images projetées. On peut inversement envisager d'autres possibilités d'organisation du dispositif qui induisent quant à elles un effacement du locuteur. Il importe alors de considérer ce que le spectateur est autorisé à voir de cet appareillage. L'occultation ou l'exhibition de la source est un facteur essentiel pour la définition de la relation horizontale, dans la mesure où il touche au caractère illusionniste de la représentation.

La dialectique de l'exhibition et de l'occultation ne s'applique donc pas qu'aux œuvres de la seconde moitié du XX<sup>e</sup> siècle qui visent à déconstruire la représentation traditionnelle. La réflexivité qu'elles postulent se manifestait déjà – certes dans une intention et avec un effet autres – dans des pratiques d'exhibition d'appareils de (re)production de la voix. Lorsqu'on coupla le phonographe à l'image animée, on s'ingénia généralement à dissimuler la

54. R.-M. Arlaud, *op. cit.*, p. 101.

55. André Gaudreault (collaboration Germain Lacasse), «Fonctions et origines du bonimenteur du cinéma des premiers temps», *Cinémas*, Vol. 4, N° 1, 1993, p. 139.

56. Jean Châteauvert, «Regards à la caméra: les comédiens en point de mire», in Michel Marie et Laurent Le Forestier (dir.), *La Firme Pathé Frères 1896-1914*, Paris: AFRHC, 2004, p. 232.

machinerie, comme cela fut plus tard le cas pour le projecteur (voir II.4). Par contre, le phonographe donna lieu à un autre type de spectacle, inscrit dans la lignée des démonstrations scientifiques et qui ne dura guère au-delà de la période où l'appareil faisait office d'attraction, c'est-à-dire lorsqu'il était exhibé individuellement et pour lui-même. Ces spectacles n'étaient pas associés à la production d'une image, mais comprenaient l'intervention d'un locuteur *live*. À l'instar de la plupart des projections sonores de l'époque du « muet », on y trouve une alternance entre la voix d'un présentateur et la voix phonographiée. Le lien entre les deux types de voix est par contre plus étroit dans la présentation du phonographe, puisqu'il arrivait fréquemment, comme l'indiquent des comptes rendus d'époque, que le locuteur énonce une phrase ensuite répétée par le pavillon : cette authentification de la réversibilité de l'appareil confère à la phrase reproduite une valeur performative très forte. L'exhibition du pôle machinique s'effectuait par contre au détriment de l'humanisation de la reproduction vocale lorsque le présentateur faisait tourner le cylindre dans le sens inverse de celui de l'enregistrement. Cette expérience, réalisée entre autres par un dénommé Puscus, représentant d'Edison en Grande-Bretagne, témoigne d'une volonté d'exploiter la valeur attractionnelle du mécanisme proprement dit, comme le feront plus tard les frères Lumière en faisant défiler par exemple la *Démolition d'un mur* (1895) à l'envers<sup>57</sup>. L'utilisation pratiquée par Puscus succédait néanmoins, comme l'explique le chroniqueur du journal scientifique *Engineering* qui assista à cette séance donnée à l'« Institution of Civil Engineers », à une présentation de la machine par elle-même explicitement placée sous le signe de l'anthropomorphisme, puisque l'appareil, selon ses dires, « prononça avec une voix encore plus claire qu'auparavant : « Comment allez-vous? », « Est-ce que le phonographe vous plaît? », puis commença à rire d'un rire profondément humain »<sup>58</sup>. Les adresses à l'auditoire placées dans la « bouche » de l'appareil inscrivent la présentation dans une logique d'humanisation, en dépit de la monstration du dispositif qui caractérise ce spectacle. Le phonographe semble fictivement se départir de sa nature mécanique, car les phrases prononcées lui permettent de se référer à lui-même, ce qui le constitue en « individu ».

57. Voir George Sadoul, *Histoire générale du cinéma*, tome 1, Paris : Denoël, 1948, p. 360.

58. « The Phonograph », *Engineering*, 8 mars 1878, disponible sur le site : [members.lycos.co.uk/MikePenney/leonscott.htm](http://members.lycos.co.uk/MikePenney/leonscott.htm)

Ce mode de présentation ainsi que le texte du journaliste, qui use du discours direct et humanise la description du rire phonographié, obéissent à une volonté d'atténuer ce que l'on pourrait appeler «l'inquiétante étrangeté» de la déliaison fondamentale: comme dans la situation décrite par Freud<sup>59</sup>, le spectateur habitué par le synchronisme audiovisuel à une régression à l'«animisme des primitifs» se voit confronté à la dissociation de l'apparence et de la voix, scission effective mais généralement refoulée dans la mesure où elle renvoie à un éclatement de l'unité du sujet qui concerne également l'audiospectateur. À l'époque de l'essor de machines à représentation, la dichotomie de l'humain et de la machine engage des phénomènes nouveaux de réception.

Une grille établie sur le modèle proposé pour les dispositifs visuels par François Albera et Maria Tortajada offre une vision synthétique de ces divers aspects. À travers l'énumération raisonnée de cas de figure, je ne vise pas à l'exhaustivité – qui serait d'autant plus utopique dans le champ fort hétérogène des pratiques *parlées* constitué d'innovations ponctuelles et éphémères dues aux exploitants – mais j'entends avancer certains axes d'analyse. De toute manière, cette typologie doit être combinée avec les diverses composantes de la machinerie nécessaire à la production de l'image, telles que les exposent Albera et Tortajada. Les types sont énumérés ci-dessous dans un ordre croissant de «dématérialisation» du locuteur, afin que le dernier cas listé pour chaque point représente celui qui offre le plus de parenté avec le procédé de la voix-over, but ultime de notre étude. Le premier axe correspond au *rapport entre le spectateur et la «machinerie»*, que je propose d'aborder en fonction des trois paramètres suivants:

#### TYPE DE SOURCE

- Voix d'un locuteur visible dans la salle, s'exprimant de vive voix;
- Voix d'un locuteur visible dans la salle, médiatisée par un instrument technique (porte-voix, microphone et amplificateur ou modulateur électrique, etc.).
- Voix d'un locuteur dissimulé.
- Voix phonographiée ou produite mécaniquement dont l'appareillage est visible.

59. Sigmund Freud, «L'inquiétante étrangeté» [1919], in *L'Inquiétante étrangeté et autres essais*, Paris: Gallimard, 1989, p. 245.



- Cet appareil peut être activé: par le spectateur (Kinétophone, poupée-phonographe), par le *showman* (spectacles phonographiques) ou de façon mécanique (systèmes de synchronisation).
- Conjonction de la voix d'un locuteur et de la voix phonographiée, sur le mode de la simultanéité ou de l'alternance.
  - Voix phonographiée dont l'appareillage est dissimulé.

#### **EMPLACEMENT SPATIAL DE LA SOURCE VOCALE**

- Lieu d'origine de l'image.
- Espace spectral.
- Espace scénique aménagé dans la salle.
- Espace situé derrière l'écran;
- Lieu impossible à déterminer pour cause d'acousmatisation, sources secondaires multiples (dispositif englobant).

#### **MODE D'ACCÈS A LA VISIBILITE DE LA SOURCE**

- Exhibition immédiate et permanente du locuteur ou de l'appareillage.
- Exhibition provisoire du locuteur ou de l'appareillage (prologue, intermède).
- Accommodation progressive au représenté audiovisuel par effacement de la source effective.
- Occultation constante de la source.

Il est évident que des liens privilégiés se créent entre certains de ces aspects. Par exemple, la dissimulation du locuteur, qui procède d'une volonté illusionniste, tend à favoriser l'espace situé derrière l'écran, puisque le support de l'image « fait écran » et attire sur lui toute l'attention des spectateurs. Du point de vue de l'ancrage de la voix dans la représentation, le degré d'exhibition de la source vocale est donc inversement proportionnel à la subordination du son à l'image. À ces caractéristiques relatives à l'organisation spatiale effective de la salle, il faut ajouter la façon dont le locuteur se réfère au lieu du spectacle: il s'agit alors de déterminer quelle est la fonction linguistique dominante du texte proféré (voir II.3).

Lorsqu'on traite conjointement du verbal et de l'iconique, l'étude du dispositif permet d'aborder la question de la « subordination » d'une composante à l'autre en dépassant la seule dimension textuelle. L'exigence d'une hiérarchisation des deux pôles de la production apparaît par exemple lors de séances procédant à un

couplage du phonographe et du cinématographe, puisque la vitesse de défilement de la pellicule peut être calquée sur celle de la rotation du cylindre ou du disque (l'inverse est aussi possible, mais non sans modifier considérablement la hauteur des sons reproduits). Il se peut que cette relation verticale propre à la machinerie intervienne également lors de la phase du tournage, certains éléments profilmiques devant se soumettre aux conditions prophonographiques (par exemple dans la situation du *play-back* sur le plateau). De toute manière, le mode de *présence* est capital : lorsqu'il s'agit d'un locuteur en chair et en os, la façon dont il s'offre au regard du public – et lui offre son regard – est déterminante : ainsi la *benshi* Midori Sawato se tenait-elle sur un podium situé entre les spectateurs et l'écran, et sa présence était soulignée grâce à un éclairage lorsqu'elle accompagna la projection du film *Orochi* (Buntaro Futagawa, 1925) lors de la rétrospective de films japonais muets dans le cadre du 20<sup>e</sup> Festival de Pordenone d'octobre 2001. G.-M. Coissac recommandait également une position surélevée pour le conférencier<sup>60</sup>. Il se peut donc que la place du bonimenteur soit mise en évidence par des éléments visuels de nature scénographique.

Par ailleurs, on peut retrouver l'opposition inhérente au dispositif entre l'humain et la machine (constituée au minimum de l'appareil de projection) à l'intérieur même de la performance du locuteur, celui-ci étant susceptible de disposer de moyens technologiques variés, à l'instar de la « machine à bruits » que le bonimenteur avait coutume d'actionner, ainsi que le note Jean A. Keim<sup>61</sup>. Le plus rudimentaire de ces instruments est probablement le porte-voix, indispensable dès lors que le spectacle est prévu pour de nombreux spectateurs. Cet attribut est plutôt lié au domaine du cirque, mais aussi à des spectacles contemporains de l'émergence de la culture de masse dans laquelle s'inscrira le cinéma. On peut penser par exemple au *Wild West Show* de Bill Codi (lancé en 1883, mais dont le succès s'est poursuivi jusqu'au début des années 10) – spectacle dépeint dans le film *Hidalgo* (Joe Johnston, 2003), où « Buffalo Bill » commente le spectacle à l'aide d'un porte-voix sur des airs joués par une fanfare militaire – dans lequel son promoteur interprétait un meneur de jeu et bonimentait oralement

60. G.-M. Coissac, *op. cit.*, p. 175.

61. J. A. Keim, *op. cit.*, p. 41.

certaines scènes. Dans ce *show*, la parole permettait de conférer une dimension mythique aux actions passablement triviales qui étaient mises en scène dans l'arène. On retrouve là l'un des rôles fondamentaux de l'oralité qui consiste à assurer la transmission des récits fondateurs. La manifestation vocale d'un narrateur a d'ailleurs été confirmée par l'analyse des programmes du spectacle de Codi entreprise par Kristen Whissel. Il ressort de cet examen que ce promoteur et son associé Salsburry avaient pour objectif de «présentifier» le spectacle grâce à des adresses, et recouraient simultanément à une représentation au fort pouvoir illusionniste. Le texte constituait un facteur d'explication et d'authentification, mais se faisait également le vecteur du caractère attractionnel de cette représentation<sup>62</sup>. Cet amalgame de voix-explication et de voix-attraction devait caractériser nombre de «conférences avec projection» du début du siècle. D'ailleurs, le *show* de Codi a probablement exercé, en s'internationalisant, une influence sur certains films de la même époque: Martin Sopocy, dans l'article où il démontre l'importance des commentateurs lors des projections des films de James Williamson, suppose que ce dernier s'est inspiré pour *Attack on a China Mission* (1900) «des *Wild West Show*, alors populaires en Europe depuis la première tournée de Buffalo Bill en Angleterre et sur le continent en 1887»<sup>63</sup>. Bien que Sopocy ne fasse aucune allusion à la fonction d'un «narrateur» (il s'agit pourtant là de l'objet de son intervention, pionnière en la matière), on est en droit de supposer que le type de manifestations vocales que comportaient les *shows* de Codi a également influencé la façon de présenter certaines projections de la même époque. Jacques Portes mentionne un autre spectacle directement inspiré de la mise en scène des mythes de l'Ouest, *Le Drame de la civilisation*, dans lequel «un narrateur raconte les épisodes et la machinerie fait souffler un cyclone»<sup>64</sup>. Ces pratiques s'apparentent fortement au mélange de boniment et d'effets sonores que l'on trouve dans le cinéma des premiers temps. Avec l'évolution technologique, le locuteur intégré à un dispositif machinique sera également en

62. Kristen Whissel, «Placing the Spectator on the Scene of History: the Battle Re-enactment at the Turn of the Century, from Buffalo Bill's Wild West to the Early Cinema», *Historical Journal of Film, Radio and Television*, Vol. 22, N° 3, 2002, p. 227.

63. Martin Sopocy, «Un cinéma avec narrateur. Les premiers films narratifs de James A. Williamson», *Cahiers de la Cinémathèque*, N° 29, 1979, p. 114.

64. Jacques Portes, *De la scène à l'écran. Naissance de la culture de masse aux États-Unis*, Paris: Belin, 1997, p. 65.

mesure d'utiliser des appareils capables d'amplifier, voire de modifier sa voix, opérant ainsi un désancrage spatial (diffusion par des enceintes distantes du locuteur) et une mécanisation de la production vocale. Ces deux aspects sont inhérents à la technique du phonographe qui, de surcroît, diffère dans le temps l'émission phonographique.

Le critère de «l'alternance» entre parole vivante et voix enregistrée implique de considérer également la situation inverse, celle de la «simultanéité». Le recouvrement de la parole vivante par une voix phonographiée (ou inversement) est surtout théorique: on imagine en effet assez mal quels avantages cette situation peut fournir dans le cadre d'une conception traditionnelle de la représentation sonore, puisqu'il conduit à une perte d'intelligibilité, surtout lorsque le texte proféré par les deux voix n'est pas identique. On ne peut toutefois exclure a priori de tels phénomènes de chevauchement comme on en trouve fréquemment dans les films de Jean-Luc Godard. En dehors de l'institution cinématographique, on trouve ce type de superposition chez les médiums charlatans, qui en usaient à des fins d'authentification. Dans la description d'une séance de spiritisme citée par Conan Doyle, un initié affirme par exemple qu'il est préférable que le médium parlât en même temps que les voix des esprits, afin que les participants de la séance soient «convaincus que les voix ne sont pas dues à une quelconque ventriloquie»<sup>65</sup>. Cette fonction d'authentification de la profération simultanée est représentée dans la fiction *Les Automates* de E. T. A. Hoffmann, puisque les deux protagonistes principaux écartent tout soupçon de ventriloquie lorsque l'artiste chargé de la présentation d'un «Turc parlant» s'adresse au public en même temps que la machine<sup>66</sup>. On peut imaginer que cette fonction d'authentification incombait également à la voix du *showman* des premières présentations du phonographe, les sons pouvant être perçus, en raison de leur nature acousmatique, comme des manifestations surnaturelles. En effet, la présence d'un individu muet à côté de l'appareil avait fait peser les mêmes doutes sur certains observateurs, même avertis. Le chroniqueur de *L'Illustration* du 23 mars 1878 rapporte par exemple:

65. Arthur Conan Doyle, *Histoire du spiritisme*, Paris: Éditions du Rocher, 1981 [1<sup>re</sup> édition 1927], p. 310.

66. E. T. A. Hoffmann, «Die Automate», in *Die Serapiens-Brüder*, tome 1, Darmstadt: Wissenschaftliche Buchgesellschaft, 1968 [1<sup>re</sup> édition 1813], pp. 329-330.

«De plus, l'opérateur reste toujours en présence de son instrument. Il n'ouvre jamais la bouche pendant la production de la voix d'écho. Ces circonstances font naître des doutes dans l'esprit de plusieurs physiciens. Dans les premiers moments, l'opérateur a été accusé d'être un habile ventriloque.»<sup>67</sup>

Pour le présentateur, rien n'est plus probant que de s'exprimer en même temps que la machine pour exhiber l'absence d'identité des deux sources.

Exception faite de ces cas marginaux, d'autres spectacles vivants peuvent donner lieu à une cacophonie par superposition de différentes voix. Certains artistes contemporains l'exploitent par exemple à des fins esthétiques dans des performances scéniques ou des installations. L'énonciation simultanée d'une même voix et d'un même texte (et non seulement d'un même contenu comme dans certaines techniques de doublage) par deux canaux différents provoquent alors des effets de *déliasion* qui déstabilisent le spectateur.

L'exhibition de la mécanicité du phonographe demande aussi à prendre en compte l'éventualité d'un rapport physique du spectateur avec le dispositif sonore. On différenciera alors un système comme le Phonorama (Berthon, Dussaud et Jaubert, 1898), où chaque spectateur devait coller l'oreille à un récepteur téléphonique pour que lui parviennent les sons accompagnant l'image projetée, du Phono-Cinéma-Théâtre de Clément-Maurice (1900), dont le son émanait d'un phonographe à cylindre situé dans la fosse d'orchestre, l'usage du combiné ne s'effectuant qu'à l'intérieur même de «l'espace de la production» pour relier le phonographe à la cabine de projection (où l'opérateur réalisait une synchronisation *live*). Le Phonorama alliait donc un dispositif visuel englobant (public collectif face à un écran) à un dispositif sonore individuel; le Kinétophone d'Edison (celui de 1895) permettait par contre un accès individuel du spectateur aux deux composantes. La manipulation individuelle est un facteur particulièrement important, car elle ne se limite pas à enclencher l'appareil, mais permet à tout moment d'atténuer ou de supprimer le son alors que l'image défile. Dans une salle de cinéma traditionnelle, l'intensité sonore est presque identique pour tous les spectateurs. Le son s'impose d'autant plus qu'on ne peut pas «fermer» les oreilles, tandis qu'il est aisé de fermer les yeux.

67. Cité par G. Pisano, *Une archéologie...*, *op. cit.*, p. 151.

L'emplacement de la source sonore dans la salle est quant à lui difficilement concevable indépendamment de ses implications sur la représentation audiovisuelle. En effet, dès les tentatives de couplage d'un phonographe et d'un appareil de projection cinématographique, le choix de placer le-s phonographe-s derrière l'écran (ou le long de celui-ci) s'est rapidement imposé – en dépit des nombreux avantages pratiques qu'aurait présentés une proximité des deux appareils – car, selon Harry Geduld, «il ne semblait pas naturel pour le public d'écouter un son provenant de derrière leur siège alors qu'ils faisaient face à l'image»<sup>68</sup>. On fit donc primer l'unicité de la représentation audiovisuelle, garante de la complétude diégétique, sur la maniabilité de la machinerie. À des fins identiques, Léon Gaumont expérimenta une technique consistant à déplacer manuellement les phonographes derrière l'écran pour correspondre aux déplacements visibles sur l'écran<sup>69</sup>. Cette adaptation de la machinerie sonore à l'espace écranique et, partant (mais non sans problème), diégétique, constitue un équivalent «matériel» des technologies stéréophoniques ultérieures. D'ailleurs, la disposition des sources sonores derrière l'écran s'est perpétuée avec la généralisation du parlant. Comme le note Rick Altman, dès que la voix est devenue, vers 1930, la «matière première» de la sonorisation au cinéma, les techniciens ajoutèrent un second haut-parleur frontal au-dessus de l'écran, réservant à la musique celui de la fosse d'orchestre<sup>70</sup>.

Ainsi que le montrent ces exemples, l'interaction entre l'image et la voix est l'une des composantes essentielles du dispositif. On peut envisager à cet égard les paramètres «verticaux» suivants:

#### **RELATIONS VERTICALES AU NIVEAU DE LA MACHINERIE**

##### *Types de projection et de source sonore*

- Images fixes/images animées/comboinaison d'images fixes et animées.
- Appareil de reproduction du son/technique de production du son/instance vocale humaine.

68. Harry M. Geduld, *The Birth of the Talkies. From Edison to Tolson*, Bloomington; London: Indiana University Press, 1975, p. 44.

69. *Ibid.*, p. 57.

70. R. Altman, *Silent Film Sound*, *op. cit.*, p. 47.

*Emplacement de la source sonore relativement à l'appareil de projection*

- Système son-image unique.
- Systèmes son-image solidaires (disjoints spatialement mais reliés mécaniquement ou électriquement).
- Système son indépendant : instance humaine pour la manipulation de l'appareil (de production ou reproduction)/ instance vocale humaine.

*Type de synchronisation*

- Absence momentanée de l'image ou du son.
- Synchronisation humaine : bonimenteur évoquant autre chose que le film/parlant du film en tant que film (présentation)/ abordant le contexte (historique, scientifique, économique, etc.) du monde du film/décrivant le monde du film/parlant à la place des personnages.
- Synchronisation humaine assistée mécaniquement, réalisée grâce à des « prothèses chronométriques » telles qu'en proposaient certains inventeurs, pionniers de l'époque « muette », aux chefs d'orchestre officiant dans les salles de cinéma, par exemple avec les systèmes nommés Visiophone, Ciné-pupitre ou Cinéorama<sup>71</sup>.
- Synchronisation phonographique : mouvement d'un cylindre ou d'un disque (également pour certaines techniques digitales) conformé au défilement de l'image (ou l'inverse).
- Synchronisation cinématographique : lecture simultanée des pistes optiques son et image sur la pellicule.

**RELATIONS VERTICALES AU NIVEAU DE LA REPRÉSENTATION**

Voix produite par la représentation visuelle elle-même (automates, dimension visuelle de la performance du bonimenteur).

Voix « synchrone » (résultat d'une synchronisation humaine, mimécanique ou mécanique).

Voix *déliée*.

Voix *off*.

Voix-over (le degré de dissociation entre l'image et la voix dépendant du mode de visualisation du locuteur présumé).

71. Voir Emmanuelle Toulet et Christian Belaygue, *Musique d'écran. L'Accompagnement musical du cinéma muet en France 1918-1995*, Paris : Réunion des Musées nationaux, 1994, pp. 76-77.

Bien que j'utilise le terme générique de «son», les éléments énumérés concernent en priorité la voix, ce qui explique l'absence de référence aux effets sonores ou à la musique. On distingue à travers ces catégories un partage entre les sources mécaniques et humaines qui prend une signification particulière dans le cas de la voix, fondamentalement associée à l'être humain. Je proposerai différents axes d'analyse de cette interaction entre l'homme et la machine au sein du dispositif, d'abord en consacrant les deux premiers chapitres au bonimenteur des vues animées, puis en examinant les relations verticales internes à la représentation. Il s'agira alors de distinguer un certain nombre de régimes vocaux et d'analyser le fonctionnement sémiotique, narratif et énonciatif des interactions audiovisuelles.

## **I.2 CONFERENCIER, BONIMENTEUR ET «PHANTASMAGORE» : LE «CINEMA» PARLE**

Les paramètres définis pour la description des dispositifs audiovisuels permettent de rendre compte d'un état donné du «cinéma», ou de modalités d'exploitation singulières. Comme l'ont proposé Albera et Tortajada lorsqu'ils disent que «ce sont là quelques exemples d'éléments machiniques, constituants du dispositif <cinéma> avant le *cinéma*...»<sup>72</sup>, il est préférable de mettre le terme «cinéma» entre guillemets lorsqu'il est question de la fin du XIX<sup>e</sup> siècle, où la diversité des techniques et des pratiques – dont les combinaisons donnent lieu à des dispositifs spécifiques – interdit de concevoir, sous peine de téléologie, ces manifestations labiles et protéiformes dans un cadre monolithique et unifiant. L'hétérogénéité du spectacle se voit par ailleurs décuplée par la dimension sonore, caractérisée par l'ajout facultatif de divers éléments (bruits, musique et voix).

Les types de manifestations vocales associées à l'image animée se distribuent globalement selon une ligne de partage définie par l'opposition de deux usages types dominants. Cette distinction ne correspond pas forcément à une périodisation historique, car les régimes vocaux ne se sont pas tant succédé que côtoyés durant les décennies qui ont précédé la généralisation du parlant. Il y a, d'une part, la voix non fixée, proférée *in vivo* dans la salle, qui exploite

72. F. Albera et M. Tortajada, *art. cit.*, p. 47.



pleinement le potentiel de la vocalité en occupant une position quelque peu intermédiaire dans le dispositif et, d'autre part, la voix phonographiée qui procède d'une totale intégration à la machinerie, et ouvre la voie à des possibilités de synchronisation qui influencent de façon décisive les relations verticales internes à la représentation. La différence entre un locuteur humain et un facteur technologique n'est pas sans incidences sur les rapports qu'entretiennent machinerie et représentation avec l'audiospectateur, ni sur l'attitude perceptive de ce dernier (régime de croyance, identification aux personnages de l'écran, sentiment d'appartenance à une collectivité, participation « physique » au spectacle, etc.). Ce rapport au public n'est toutefois pas déterminé par les seules conditions d'émission de la voix, mais également par les stratégies dont use le locuteur pour définir le mode d'« intégration » du public. En effet, si le partage entre voix vive et voix enregistrée correspond globalement à la distinction qui sera postulée dans cette étude entre la voix-attraction et la voix-narration, l'une et l'autre peuvent donner lieu, de façon dominante, partielle ou ponctuelle, à l'un de ces deux régimes.

En dépit des recoupements, des variantes hybrides et des développements parallèles auxquels donnent lieu les dispositifs audiovisuels, un découpage historique sommaire me semble possible si l'on considère les usages dominants de la voix durant la période sur laquelle je mettrai l'accent (1895-1935), selon des critères que l'analyse plus détaillée des discours et des productions viendra confirmer. Les années 1895-1914 se caractérisent par une présence importante (quoique facultative) de la voix sous la forme du boniment qui, dans une première phase (1895-1908 environ), se fait l'héritière de spectacles scéniques dans lesquels le cinéma se trouvait fréquemment intégré, puis, dans un second temps où s'affiche un souci de complexification narrative de la part des producteurs, se met au service de la narration filmique. Parallèlement à cette seconde phase se développe la diffusion de films avec enregistrement synchronisé, qui obéissent majoritairement à la logique attractionnelle de la première phase (la technologie du son fait office d'innovation). Ces productions minoritaires exhibent alors le « synchronisme », élément clé des futures pratiques dominantes du parlant. Si la parole, non seulement sous la forme des intertitres mais aussi des voix dans la salle, a indubitablement servi de support (certains diraient de « béquille ») à la mise en place de la continuité

narrative qui s'impose avec Griffith, elle est ensuite évincée au profit d'une autonomisation du discours visuel qui profite aux tentatives de légitimation artistique du cinéma (muet) de la seconde moitié des années 10. La Grande Guerre correspond donc approximativement à la disparition du cinéma « parlé », qui ne se perpétue plus dès lors que sous la forme de présentations « péri-filmiques » (prologues, épilogues ou intermèdes). En effet, contrairement à l'image animée, la parole est commune à d'autres moyens d'expression fortement implantés culturellement (théâtre, littérature) dont le cinéma devait se distinguer pour ne pas leur paraître subordonné. Pour les partisans du « 7<sup>e</sup> Art », le médium se devait de produire du sens par ses moyens propres, que l'on réduisait aux seules caractéristiques de l'image. Ce contexte de réception explique que la musique d'accompagnement n'ait pas été évaluée selon les mêmes critères que la voix, puisque critiques et théoriciens ne considéraient pas celle-là comme un élément susceptible de nuire au caractère visuel du spectacle, mais comme un complément de la « musique des images ». Alors que la parole fait appel à l'intellect, la musique s'adresse, à l'instar du flux d'images auquel elle se mêle, au registre des sensations dont elle amplifie l'impact par ses effets « synesthésiques ». Ainsi, par exemple, Ricciotto Canudo, l'un des principaux partisans d'une conception artistique du médium, rapproche dans un texte de 1921 la musique du cinéma parce qu'elle est capable, « par l'expression infinie des états d'âme, par son évocation large des grands sentiments élémentaires *sans la précision rigide des pensées et des mots* », d'atteindre au statut de « langage universel »<sup>73</sup>. Dans un article rédigé deux mois plus tard, Canudo parle certes des « chansons filmées » de Lordier comme d'une avancée sur le terrain des expérimentations de synchronisation, mais ne manque pas de préciser qu'il serait nécessaire d'en évacuer la composante verbale<sup>74</sup>. Durant la décennie 1917-1927, la parole n'est donc plus à l'honneur ; c'est pourquoi le travail d'inventeurs comme Lordier ne réussit pas à exercer une influence sur les pratiques dominantes du « muet ». Dans son étude des discours émis par divers champs au cours des deux décennies 1910-1920 en France, Laurent Guido démontre avec force sources que la notion de « rythme » constituait un interprétant central des

73. Ricciotto Canudo, *L'Usine aux images*, Biarritz : Éditions Séguiet, 1995 [1<sup>re</sup> édition 1927], p. 72.

74. *Ibid.*, p. 92.

réflexions sur le cinéma de cette époque – le film *La Roue* de Gance est par exemple encensé par Dulac parce qu'il relève d'une «symphonie visuelle» – l'intérêt s'étant déplacé vers des préoccupations musicalistes qui s'accompagnèrent d'une mise à l'écart du langage verbal<sup>75</sup>. La parole réapparaît par contre sur le devant de la scène avec la généralisation du parlant, d'abord sous une forme de mise en film du chant qui n'est pas sans parenté avec certaines Phonoscènes Gaumont des années 1906-1914, puis grâce à l'incontestable domination d'un certain type de voix synchrone, asservie à la transmission d'informations narratives. Avant que les situations dialogiques ne mènent à une importante standardisation des pratiques, nous assistons à une période d'expérimentations. Le caractère novateur des solutions proposées durant l'inter règne s'explique autant par l'imperfection des techniques que par le bouillonnement théorique suscité par un contexte nouveau où la voix du cinéma «cherche sa voie».

Ce panorama succinct serait probablement contredit par nombre d'analyses de détail, car on peut bien sûr trouver des exceptions parmi les diverses recherches des pionniers de la sonorisation au cinéma, travaux souvent isolés et au succès éphémère. D'ailleurs, si l'on prend cette périodisation à la lettre, cela laisserait entendre qu'il existe une véritable universalisation des principes régissant les pratiques qui allient son et image, alors qu'il est nécessaire, ainsi que l'a montré Rick Altman, de respecter la complexité et la diversité des phénomènes<sup>76</sup>. Par exemple, il est établi que le cinéma parlé s'est prolongé au Japon jusqu'à la fin de ce que fut en Europe l'«inter règne du parlant», les performances du *benshi* étant intégrées au nouveau dispositif (alors que la production nationale était encore muette). Toutefois, le cadre général que j'ai esquissé pour le passage du parlé au parlant vaut surtout pour la France et les États-Unis, deux pays sur lesquels il me semble légitime de se concentrer, non seulement parce que c'est là que les recherches sur le son ont été le plus amplement développées et théorisées, mais aussi parce qu'il s'agit d'importantes nations productrices de films dont l'influence fut considérable. Pour la période des premiers temps,

75. Laurent Guido, *L'Âge du rythme*, thèse de doctorat de l'Université de Lausanne, 2004. Voir également Dominique Chateau, «Le rôle de la musique dans la définition du cinéma comme art: à propos de l'avant-garde des années 20», *Cinémas*, Vol. 3, N° 1, 1992. Chateau indique dans son introduction qu'il existe à cette époque une partition entre les arts qui sont valorisés du point de vue de l'analogie qu'ils présentent avec le cinéma, et ceux qui, basés sur le verbe, sont écartés (*ibid.*, p. 80).

76. R. Altman, *Silent Film Sound*, *op. cit.*, pp. 9 et 13.

Edison et Gaumont font figure de pionniers dans le domaine du couplage du son et de l'image; plus tard, la généralisation du parlant est orchestrée par les *majors* hollywoodiennes (alliées en Europe avec l'Allemagne) sur la base de technologies qui, elles, sont le fruit de recherches menées dans différents pays.

Dans le cadre de la dichotomie du parlé et du parlant, la voix s'inscrit dans le paradigme oppositionnel de l'humain et de la machine. À cet égard, le boniment présente une conjonction intéressante des deux termes en exhibant le pôle humain. C'est pourquoi il me paraît important d'élargir la question de la « voix dans les films » aux pratiques vocales *live* qui, lorsqu'elles intègrent (ou sont intégrées à) des projections animées à l'époque du cinéma des premiers temps, donnent lieu au « cinéma » *parlé*.

Depuis que le bonimenteur a été constitué en objet d'étude par les historiens du cinéma, la comparaison du « cinéma » parlé avec la voix-over fait presque office de lieu commun. Toutefois, si l'on se contente de mentionner le bonimenteur à titre d'« ancêtre », on risque de sous-estimer les spécificités du dispositif auquel il appartient, ainsi que les conséquences de l'exploitation de la voix vive. Inversement, la voix phonographiée du parlant transforme fondamentalement les conditions de production et de réception de la parole proférée. Après avoir examiné en quoi le spectacle de vues animées bonimenté est le lieu d'exercice d'un régime vocal singulier, celui de la voix-attraction, j'envisagerai sous quelles formes ce régime fait retour à l'époque du parlant.

La distinction entre parlé et parlant a trait à la nature du dispositif. Le *cinéma parlant* renvoie à une pratique institutionnalisée d'asservissement du verbal au monde du film: on (se) parle *dans* le film. Le participe présent « parlant » (« talking ») suggère que le film occupe une position de *sujet* (il est un *talkie*), ce qui révèle une conception verbocentrée. La catégorie du « cinéma » *parlé* recouvre quant à elle une série de pratiques hétérogènes qui ne s'inscrivent pas dans la tradition phonographique, mais ressortissent aux spectacles vivants. Le film *parlé* est considéré comme un *objet* au sein d'un dispositif qui, pris dans son ensemble, est parlant. Dans le film *parlé*, on (me) parle *du* film. Cette question a déjà été abordée succinctement par certains théoriciens. Michel Chion propose par exemple la notion de « parole-écran » pour qualifier une émission vocale qui, issue du hors-champ ou de l'espace « off », interpelle le

personnage ou commente ses actions<sup>77</sup>. François Niney souligne l'indécision dans laquelle demeure l'origine de la parole dans un film documentaire en définissant ainsi la «voix On»: «On: on nous parle sur les images, on nous dit ce qu'il faut en penser. On, c'est le Médium...»<sup>78</sup> Or, même lorsque le commentateur *over* occupe manifestement une position de retrait par rapport aux images, ce «on» du cinéma parlant n'a pas d'équivalent dans le cinéma bonimenté, où le locuteur n'est pas assimilé par l'audio-spectateur à la source «impersonnelle» de la machinerie, mais représente indéfectiblement un «je» présent dans l'espace de la salle. La parole du «cinéma» *parlé*, extérieure sur le plan matériel à la diégèse filmique, ne peut être appréhendée qu'en tant que composante du dispositif «cinéma». Par ailleurs, ce ne sont pas seulement les éléments appartenant au monde de l'image projetée qui sont visés par la voix du «cinéma» *parlé*, mais c'est aussi le film lui-même. Le texte proféré peut donc endosser une fonction déictique (en se référant à *ce* film) relativement à une *deixis* réelle qui s'origine dans le lieu où le spectacle est proposé.

Précisons que les modalités spectaculaires du «parlant» et du «parlé» ne coïncident pas forcément avec l'opposition «mécanique» *vs.* «humain», puisque le parlé peut recourir accessoirement à des appareils de production ou de restitution des sons. De telles séances demeurent néanmoins régies par les conditions du «cinéma» parlé. Inversement, certains films pouvaient «être tournés dans l'attente expresse d'une telle exploitation parlée»<sup>79</sup>, si bien que la situation d'oralité influençait en amont la représentation qui advient grâce à la machinerie.

La généralisation du cinéma dit «parlant» ne fut pas le premier indice du statut particulier que l'on accorde à la voix humaine au sein des productions sonores et visuelles. Durant la période du «muet» qui précède la légitimation artistique du médium, l'absence de la voix était considérée par certains (l'un des textes les plus célèbres à cet égard est celui de Maxime Gorki)<sup>80</sup> comme le manque essentiel de la représentation cinématographique.

77. Michel Chion, *Un art sonore, le cinéma*, Paris: Éditions des Cahiers du cinéma, 2003, p. 428.

78. François Niney, *L'Épreuve du réel à l'écran*, Bruxelles: De Boeck, 2000, p. 136.

79. R. Altman, «Naissance de la réception classique. La campagne pour standardiser le son», *Cinémathèque*, N° 6, 1994, p. 102.

80. Maxime Gorki, «Le cinématographe Lumière», 6 juillet 1896 [trad. française de Valérie Pozner], 1895, N° 50, décembre 2006.

En outre, la voix a connu de nombreuses modalités d'intégration au spectacle audiovisuel durant le « pré-cinéma » et le « cinéma des premiers temps ». Avant d'en venir à la figure du bonimenteur qui a acquis une place centrale dans l'étude historique de cette période, il me semble utile de rappeler quelques dispositifs « audiovisuels » qui, avant le « cinéma » parlé, comprenaient un locuteur. Ces usages de la voix contribuèrent probablement à définir de façon décisive certaines modalités de l'exploitation cinématographique des premiers temps.

Les recherches menées en archéologie du cinéma depuis quelques années par des historiens comme Jacques Perriault et Laurent Mannoni ont permis d'exhumer plusieurs pratiques spectaculaires qui convoquaient, bien avant l'apparition des projections cinématographiques bonimentées, la présence conjointe d'une voix et d'un système optique<sup>81</sup>. Or les conférences de lanterne magique ont contribué, à l'instar d'autres pratiques, à fixer un horizon d'attente auprès des (futurs) spectateurs du cinématographe, et à instaurer des habitudes parmi les promoteurs. Comme le note Rick Altman, le « cinéma » connut à ses débuts diverses définitions en fonction du spectacle plus ancien auquel il était associé; ainsi, « lorsque les films étaient considérés comme des plaques de lanterne, ils semblaient exiger un conférencier vivant situé à côté de l'écran et expliquant chaque image »<sup>82</sup>. Par exemple, on peut dire que le discours démystificateur tenu par Athanase Kircher dans son *Ars magna* (1646) à propos de son système catoptrique modela historiquement le « spectateur » dans son rapport à la machinerie en favorisant un certain type de relation horizontale<sup>83</sup>. Au vu de la démarche pédagogique prônée par Kircher, le discours verbal devait primer sur la machinerie et sur la représentation visuelle. Le potentiel attractionnel de l'une et de l'autre était toutefois exploité, et rien n'interdit que la voix pût exprimer l'engouement envers ce que les plaques représentaient, voire à propos du moyen de représentation proprement dit. D'ailleurs, la *camera obscura*, première exploitation de principes

81. J. Perriault, *op. cit.*; Laurent Mannoni, *Le Grand Art de la lumière et de l'ombre*, Paris: Nathan, 1995.

82. R. Altman, *Silent Film Sound*, *op. cit.*, p. 126.

83. Voir Charles Musser, *The Emergence of Cinema: The American Screen to 1907*, New York; Toronto: Scribner's Sons; Collier Marmillon, 1990, p. 18; L. Mannoni, *Le Grand Art...*, *op. cit.*, p. 33.

optiques à avoir permis une mise en spectacle de l'image « projetée » (ou plutôt « reflétée »), connu à travers ses développements une grande diversité d'usages (enseignement religieux ou scientifique, diablerie, satire, récits historiques, etc.). Ces champs d'application se déploieront dès la popularisation des connaissances sur le fonctionnement de l'appareil, notamment transmises par le truchement des éditions successives de l'ouvrage *Magiae naturalis* (1558) de Giambattista Della Porta, et de la tournée européenne du spectacle de lanterne magique de Thomas Walgensten.

Della Porta conçoit le principe de la chambre noire comme un véritable spectacle, tant au niveau de l'image (il prévoit d'organiser des mises en scène de l'autre côté du mur perforé) que du son (il propose d'accompagner de musique la monstration d'images). On peut imaginer que le dispositif de Kircher donna également lieu à des variantes auprès de promoteurs moins soucieux des vertus éducatives de l'appareil. Il est en effet vraisemblable que la *camera obscura*, la lampe catoptrique ou cette lanterne justement nommée « magique » furent des instruments de leurre aux mains de charlatans. C'est précisément dans ce contexte où la visée première n'est pas (seulement) de divertir, mais de susciter une croyance dans l'existence effective de ce qui n'est en fait qu'une représentation, qu'intervient le pouvoir évocateur de la voix. Le moine parisien Jean-François Nicéron réagit en érudit face à ces spectacles capables de duper ceux qui ignorent les principes d'optique. Il décrit ainsi ces projections qu'il nomme « Perspectives » :

« S'il y a quelqu'un caché derrière la charte qui fasse l'esprit, comme l'on dit, en parlant comme ceux qui font danser les marionnettes, les simples croiront que ce sont les personnes du tableau qui parlent, car on leur voit ouvrir la bouche et remuer les lèvres. »<sup>84</sup>

La voix du locuteur est dissimulée derrière l'écran, sur le modèle du spectacle de marionnettes, auquel Nicéron compare ce dispositif. Le pouvoir illusionniste est cependant supérieur dans ces « Perspectives » puisqu'elles proposent la transposition spatiale d'un fragment de réalité dont l'existence est effective. Nicéron utilise l'expression « faire l'esprit », qui suggère l'exploitation d'une voix

84. Cité par L. Mannoni, *Le Grand Art...*, *op. cit.*, p. 21.

déliée des corps représentés; il ne s'agit toutefois pas, comme le décrit Mannoni<sup>85</sup>, d'une «voix-off» (que l'on comprenne le terme «off» au sens de «hors-champ» ou de «over»), à moins que nous nous plaçons du point de vue des producteurs du spectacle audiovisuel. Car, pour les «audiospectateurs» de la chambre noire, il s'agit, du moins dans l'idéal, d'une représentation synchronisée. Cette technique élémentaire, héritière de pratiques scéniques, préfigure les accompagnements verbaux du cinéma muet réalisés par des acteurs cachés derrière l'écran, comme en pratiquèrent, au cours des années 10<sup>86</sup>, Georges Lordier en France<sup>87</sup>, les «ciné-déclamateurs» en Russie<sup>88</sup> et de nombreuses compagnies américaines (Humanovo, Actologue, Humanophone, Natural Voice Talking Pictures, etc.) inscrites dans des circuits dont l'organisation était calquée sur celle des vaudevilles<sup>89</sup>. Dans ce type de pratiques, la voix provient du même espace que celui où apparaît l'image, ce qui provoque un effet de proximité entre les composantes auditives et visuelles. Notons, pour rappeler la pérennité de l'illusion du synchronisme audiovisuel, que ce principe régit encore aujourd'hui la disposition des enceintes principales dans les salles de cinéma, placées derrière ou au bas de l'écran.

En associant des vues à un commentaire proféré par un conférencier, le spectacle de lanterne magique représente l'«ancêtre» le plus évident des projections du «cinéma» parlé. L'antériorité de cette invention technique n'exclut pas que les deux appareils aient connu une exploitation simultanée: comme le note Perriault à propos de l'institution scolaire, «l'ère de la lanterne magique et l'ère du cinéma se chevauchent et ne sont pas comme on le croit

85. *Ibid.*, p. 23.

86. Il va de soi que cette pratique eut cours dès les débuts du cinématographe, mais c'est dans les années 10 qu'elle connut une certaine institutionnalisation. Parmi les occurrences antérieures, il faut mentionner les projections organisées à Paris dans l'espace monumental de l'Hippodrome de la place Clichy entre décembre 1907 et juin 1909, où un chœur composé de plus de 100 exécutants donne «l'illusion d'entendre parler et chanter les interprètes» (Jean-Jacques Meusy, *Paris-Palaces, le temps des cinémas (1894-1918)*, Paris: CNRS, 1995, p. 174). En dépit du nombre de chanteurs, on peut s'interroger sur l'audibilité de paroles non amplifiées électriquement dans une salle comprenant 3500 spectateurs. On constate toutefois la volonté des promoteurs d'associer le réalisme de la voix synchrone au grand spectacle.

87. Voir Giusy Pisano, «Les spectacles mixtes: tradition ou anachronisme? Survivances sonores et visuelles de Robertson à Georges Lordier», in Giusy Pisano et Valérie Pozner (dir.), *Le Muet à la parole, op. cit.*

88. Voir Valérie Pozner, «La ciné-déclamation en Russie», in Giusy Pisano et Valérie Pozner (dir.), *Le Muet à la parole, op. cit.*

89. Voir Rick Altman, «Introduction: Four and a Half Film Fallacies», in *Sound Theory, Sound Practice, op. cit.*, p. 36, et «Naissance de la réception classique...», *art. cit.*, p. 102.



trop souvent strictement consécutives»<sup>90</sup>. Il s'agit en fait de deux «séries culturelles» qui se sont croisées synchroniquement<sup>91</sup>. On peut même dire que ces deux spectacles ont formé une seule et même série, celle des «vues», terme sous lequel les rédacteurs de programmes destinés aux nickelodéons regroupaient autant les plaques pour lanterne, les chansons illustrées que les bandes cinématographiques<sup>92</sup>. En ce qui concerne les techniques utilisées, Rick Altman donne un exemple symptomatique de cette réception du cinématographe comme une *extension* de la lanterne: il s'agit d'un appareil «bicéphale» («The Edison Projection Kinetoscope») comprenant deux dispositifs amovibles et interchangeables, l'un pour la présentation de plaques, l'autre pour faire défiler la pellicule<sup>93</sup>.

Le rapprochement entre les deux types de projection va néanmoins au-delà de la seule dimension technique. On a vu que certains usages de la lanterne à visée spectaculaire ont nécessairement conditionné les attentes des futurs (audio)spectateurs du cinématographe. Cette influence s'est également exercée sur le domaine de l'exploitation, les deux appareils s'adressant à un public collectif. Certains montreurs de lanterne se convertirent d'ailleurs au «cinéma» dans des pays où l'implantation de la lanterne était importante, comme en Grande-Bretagne. Lorsqu'il rappelle que James A. Williamson, le célèbre pionnier de «l'école de Brighton», était originellement montreur de lanterne magique, Martin Sopocy explique que «les organisateurs de projection, de même

90. J. Perriault, *op. cit.*, p. 117.

91. À propos de la notion de «série», je renvoie aux recherches entreprises par Édouard Arnoldy sur les deux périodes dont il est aussi majoritairement question ici, les premiers temps et le passage au parlant. Dans le champ du cinéma, le terme de «série» a été préalablement utilisé par André Gaudreault («Les vues cinématographiques selon Georges Méliès, ou: comment Mitry et Sadoul avaient peut-être raison d'avoir tort (même si c'est surtout Deslandes qu'il faut lire et relire)...», in Jacques Malthête et Michel Marie (dir.), *Georges Méliès, l'illusionniste fin de siècle?* Paris: Sorbonne Nouvelle, 1997), qui l'applique au cinéma des premiers temps afin d'accorder une attention particulière aux diverses pratiques concomitantes mais distinctes dont il procède. Édouard Arnoldy propose une définition de cette notion en précisant que «la *série* n'est pas une période, un moment circonscrit dans le temps, mais plutôt un espace *ouvert, mobile*, articulé autour d'une problématique historique, elle-même perméable aux mouvements incessants d'autres *séries*» («De la forme sérielle. L'entre du muet et du parlant: spectacle, chanson, cinéma et Cie», in Anna Antonini (dir.), *Il film e i suoi multipli*, Udine: Università degli studi di Udine, 2003, p. 327). D'un point de vue méthodologique, cette conception historique permet de ne pas rester prisonnier d'un «découpage» temporel (la «périodisation» des faits), puisqu'une série peut dériver d'une autre tout en existant simultanément à cette dernière et en suivant une évolution parallèle.

92. Rick Altman, «Technologie et textualité de l'intermédialité», *Sociétés & représentations*, N° 9, 2000, p. 12.

93. *Ibid.*, pp. 14-15.

que leur public, n'y virent qu'une nouvelle facette du spectacle de lanterne magique»<sup>94</sup>.

Durant la période du «cinéma des premiers temps» – et pas seulement dans des contextes strictement pédagogiques comme les projections réalisées entre 1902 et 1911 à Bâle par l'abbé Joye<sup>95</sup> – il n'était donc pas rare que lanterne magique et projecteur de vues animées soient utilisés en alternance, sans forcément que l'exploitant ne disposât d'un appareil multifonctions comme le projecteur Edison évoqué ci-dessus. Ce principe subsiste d'ailleurs dans les salles actuelles sous la forme de diapositives projetées, pratique qui obéit néanmoins, à une ère où le cinéma est clairement défini comme média autonome, à un partage plus net entre images fixes et animées, la très grande majorité des diapositives se voyant cantonnées à une fonction publicitaire et à un mode de présentation qui diffère peu de celui des affiches murales. Avant que l'installation d'un second projecteur de vues animées ne se généralise durant les années 10, les exploitants disposaient habituellement d'une lanterne unique qu'ils utilisaient aussi pour des spectacles spécifiques (les «chansons illustrées» aux États-Unis) ou pour s'adresser aux spectateurs. Dans ce dernier cas, les annonces entretiennent, à l'instar des diapositives actuelles, un rapport d'extériorité totale à la diégèse créée par la représentation, et concernent plus directement l'espace même de la salle. Parmi les nombreuses plaques de ce genre citées par Altman, on en trouve une dont le texte enjoint le public à ne pas se retourner<sup>96</sup>. Cette image dessinée représente des spectateurs de cinéma qui détournent la tête en direction du projecteur (rejeté presque intégralement hors-champ, comme si on voulait en évacuer la présence) alors que le texte «Une autre image dans une minute, l'opérateur est en train de changer le film» apparaît sur l'écran. Une telle mise en abyme du dispositif montre que l'usage de la lanterne ne consistait pas seulement en une intégration mutuelle des images animées et fixes, mais fonctionnait également sur le mode de l'adresse.

Les parentés éventuelles entre les plaques de lanterne et les vues cinématographiques ne doivent pas occulter une différence importante qui a trait à la question de la temporalité: l'illusion d'un

94. M. Sopocy, *art. cit.*, p. 110.

95. Voir Roland Cosandey, «L'abbé Joye, une collection, une pratique», in Roland Cosandey, André Gaudreault et Tom Gunning (dir.), *Une invention du diable? Cinéma des premiers temps et religion*, Sainte-Foy; Lausanne: Université Laval; Payot, 1992, p. 61.

96. R. Altman, *Silent Film Sound, op. cit.*, p. 186.

mouvement continu accapare l'attention spectatorielle avec plus de force que l'image fixe – certaines lanternes comportaient, il est vrai, des mécanismes d'animation de l'image, mais ces procédés devaient être considérés comme une composante facultative dans une représentation globalement associée à l'image fixe – non seulement parce que cette illusion constituait une nouveauté au début du siècle, mais parce qu'elle impose un autre type de lecture visuelle. En effet, en cas de commentaire simultané à l'image, le spectateur de la lanterne pouvait à loisir porter son regard sur le conférencier, et celui-ci attirer l'attention sur sa présence ou sa performance verbale en retardant l'insertion de la plaque suivante. Ce que la narratologie filmique a parfois appelé, en reprenant la dichotomie genettienne, le « temps du récit » n'acquiert un sens en dehors la relation dialectique avec son antonyme (le « temps de l'histoire ») que dans le cas d'un médium comme le cinéma, où la durée de la projection est strictement déterminée par la longueur de la pellicule et par la vitesse de défilement – Genette utilise d'ailleurs la notion de « pseudo-temps » pour parler de la littérature, où le temps de la lecture peut être soumis à d'importantes variations individuelles<sup>97</sup>. Précisons que durant les premières années du cinématographe où l'ingérence des exploitants était plus importante qu'elle ne le sera après l'institutionnalisation de la narration (puis celle du cinéma parlant), certains projectionnistes se permettaient de faire varier la vitesse du défilement en fonction de l'action<sup>98</sup>. Dans un spectacle de lanterne magique, le réglage de la temporalité incombe par contre prioritairement au locuteur.

Ce pouvoir se manifeste ouvertement lorsque le conférencier enjoint l'opérateur à passer à la plaque suivante. Martin Sopocy affirme que « l'histoire était faite pour être *écoutée* et les images ne servaient qu'à illustrer ce récit verbal »<sup>99</sup>. Le rapport hiérarchique entre parole et image semble donc être différent de celui qui primait lors des projections cinématographiques, puisque celles-ci étaient *facultativement* accompagnées d'un boniment. Par ailleurs, le texte du commentaire proféré durant les projections fixes était publié sous forme écrite par les fournisseurs de plaques<sup>100</sup>, ce qui

97. Gérard Genette, *Figures III*, Paris: Seuil, 1972, p. 78.

98. Voir Ansje van Beusekom, « The Rise and Fall of the Lecturer as Entertainer in the Netherlands », *Iris*, N° 22, 1996, p. 131.

99. M. Sopocy, *art. cit.*, p. 113.

100. Voir Richard Crangle, « Next Slide Please: The Lantern Lecture in Britain, 1890-1910 », in Richard Abel et Rick Altman (dir.), *The Sounds of Early Cinema*, Bloomington; Indianapolis: Indiana University Press, 2001, p. 39.

constitue un indice de la vocation pédagogique du médium et de l'importance accordée par l'instance de production au contrôle exercé sur la performance orale.

L'inféodation de l'image au texte est liée à la catégorie discursive globale dans laquelle s'inscrit la projection plus qu'au type de machinerie mise en œuvre. Ainsi le discours explicatif représentait-il vraisemblablement le mode dominant lorsque les vues étaient mises au service de l'enseignement religieux. Roland Cosandey remarque que les projections animées de l'abbé Joye ne se distinguaient pas fondamentalement des projections publiques courantes avec bonimenteur, si ce n'est en ce qui concerne la subordination de l'image au verbal qu'impliquait l'orientation bien définie de ses « conférences » ; l'image, « si elle était une fin dans les conditions usuelles du spectacle, devient ici un moyen »<sup>101</sup>. G.-M. Coissac souligne combien la relation texte-image dépend du type de pratique en faisant une distinction entre le montreur de lanterne magique qui « se contente de faire défiler des vues en nommant le sujet qu'elles représentent », et le conférencier qui « traite [...] un sujet bien déterminé, pour la clarté et l'intelligence duquel il présente à ses auditeurs un certain nombre de tableaux lumineux »<sup>102</sup>. Dans son ouvrage intitulé *Manuel pratique du conférencier projectionniste*, Coissac nous livre de précieuses informations sur les spectacles lanternistes qui eurent cours à la fin du XIX<sup>e</sup> siècle et au début du XX<sup>e</sup>. Paru en 1908, ce livre de petit format constitue un abrégé des deux volumes de *La Théorie et pratique de la projection* publiés deux ans auparavant chez le même éditeur, la Maison de la Bonne Presse. Dès 1896, cet organe catholique avait confié à Coissac la gestion d'un service de projections pour lanternes, avant de le nommer en 1903 rédacteur en chef du *Fascinateur*, journal destiné aux « projectionnistes-conférenciers »<sup>103</sup>. La Bonne Presse représente à l'époque une instance d'importance en ce qui concerne la vente d'appareils de projection, de plaques et d'accessoires. Dans son manuel, Coissac ne manque d'ailleurs pas de vanter les mérites des produits estampillés de cette marque (bien visible sur tous les objets présentés dans les illustrations), qu'il s'agisse des boîtes de flacons de couleurs pour plaques<sup>104</sup> ou de tissus spéciaux pour nettoyage des

101. R. Cosandey, *art. cit.*, p. 64.

102. G.-M. Coissac, *op. cit.*, p. 180.

103. Selon L. Mannoni, *op. cit.*, pp. 5-6.

104. G.-M. Coissac, *op. cit.*, p. 126.

lentilles<sup>105</sup>. Fort de sa longue pratique de projectionniste et de plusieurs expériences comparatives menées «en laboratoire» (notamment pour tester les appareils par photométrie), Coissac est en mesure de fournir à son lectorat des recommandations portant sur tous les aspects d'une séance, tant sur le plan des manipulations techniques que sur celui du travail de l'orateur. Bien qu'il évoque par moments la projection de vues animées, la lanterne constitue à ses yeux l'instrument privilégié de la conférence, dont le but consiste à «instruire en amusant»<sup>106</sup>. Alors que les images fixes se prêtent tout à fait à la vocation didactique qu'il assigne au conférencier (ce n'est pas un hasard s'il n'utilise jamais le terme «bonimenteur», plus proche du «montreur» dont il tient à se distinguer), le cinématographe appartient selon lui à la sphère plus populaire de l'attraction foraine, entièrement vouée au divertissement :

«L'objet principal actuel de la projection, [...], c'est incontestablement le concours qu'elle fournit à l'enseignement, à la démonstration. Or, seule, la projection fixe, par l'étude du détail, la fermeté du dessin et la reproduction exacte, rend ce concours possible. Le cinématographe restera dans le domaine du barnum, il sera pour le conférencier projectionniste un moyen de récompense pour son auditoire, un élément extraordinaire, mais passager, de succès, il ne doit être et ne sera que cela.»<sup>107</sup>

En 1908, le futur auteur de l'*Histoire du cinématographe* considère l'usage de l'appareil des frères Lumière comme une attraction dont il ne faut pas abuser, comme un instrument asservi à la *captatio benevolentiae*. Soucieux de gérer l'organisation temporelle du spectacle, Coissac préconise d'agrémenter la conférence de phases récréatives basées sur la projection animée, utilisée à cette seule fin. Le cinématographe intervient comme un supplément, comme un appareil annexe – quoique parfois couplé à la lanterne, Coissac prenant l'exemple d'un modèle vendu par la Bonne Presse qui «permet de passer instantanément de la vue animée à la vue stable, et *vice versa*»<sup>108</sup> – dont le rôle se borne à éveiller la curiosité du

105. *Ibid.*, p. 41.

106. *Ibid.*, p. 187.

107. *Ibid.*, pp. 22-23.

108. *Ibid.*, p. 21.

public ou à relancer son intérêt. Coissac aborde significativement ce type de projection en mentionnant d'entrée de jeu un second appareil: le cinématographe donne «l'illusion complète du mouvement, surtout quand on lui adjoint un phonographe synchronique»<sup>109</sup>. Il peut sembler étonnant que la notion de *mouvement* soit associée à la reproduction mécanique des sons. Ce rapprochement s'explique chez Coissac par le fait que, comme le cinématographe, le phonographe (qu'il réserve visiblement à la production de bruits réalistes et non à la reproduction vocale, concurrent potentiel de l'insurpassable conférencier) contribue à accroître le pouvoir illusionniste de la représentation. Or pour lui cette fonction est subsidiaire, la conférence se définissant par son rôle pédagogique que seule la voix humaine peut endosser pleinement. Ce point ne demeure d'ailleurs pas implicite, mais il est discuté dans le *Manuel*. Ainsi, l'inconvénient du défilement continu de la pellicule cinématographique tient-il pour Coissac au fait qu'«il est bien difficile d'accompagner d'une explication intéressante, instructive, complète, les multiples phases de la scène projetée»<sup>110</sup>, c'est-à-dire à la dimension coercitive du «temps du récit» imposé à la parole vive par chaque bande. Cet argument informe la façon dont Coissac conçoit les rapports entre la parole et l'image, la première étant bien sûr privilégiée en vertu de ses aptitudes à transmettre un savoir. Tous ses conseils postulent indiscutablement une subordination de l'image au verbe, qu'il promet d'ailleurs ouvertement:

«La projection soutient la conférence, mais ne la crée pas. [...] La plupart du temps, sans la parole explicative du conférencier, l'image serait mal vue et mal comprise. En résumé, la projection est à la conférence ce que la gravure, l'illustration est au livre.»<sup>111</sup>

La série culturelle à laquelle Coissac compare le boniment est celle du livre illustré, où l'image intervient dans un second temps pour supporter un texte autonome. Coissac se réclame de cette référence écrite, répétant plus loin que «les vues viennent ainsi au secours de la parole: c'est la gravure dans le texte»<sup>112</sup>. Les fonctions respectives de l'image projetée et de la parole du conférencier

109. *Ibid.*, p. 19.

110. *Ibid.*, p. 22.

111. *Ibid.*, p. 177.

112. *Ibid.*, pp. 178-179.

conditionnent les circonstances pratiques de la séance sur un point essentiel qui, dans la réflexion de l'auteur du *Manuel*, dépend du moment que le conférencier choisit pour parler : il s'agit de la présence de la lumière. En effet, Coissac déconseille un commentaire systématiquement proféré en même temps que les images dans la mesure où l'orateur demeure constamment dans l'obscurité<sup>113</sup>. Il prône par contre une introduction verbale effectuée en pleine clarté, suivie d'un discours simultané proféré avec les lumières baissées. La visualisation du locuteur constitue par conséquent un corollaire de l'inféodation de l'image au verbe.

Cette hiérarchisation entre projection et parole, attachée à une pratique lanterniste conçue dans une intention de didactisme (voire de prosélytisme), a pu se perpétuer dans le cadre de certaines projections cinématographiques. Germain Lacasse affirme par exemple à propos des premières projections de « Passions » filmées au Québec – un « genre » filmique qui s'inscrit fortement dans la tradition lanterniste – que « la Passion était plus une conférence avec projection qu'une véritable projection »<sup>114</sup>. Isabelle Raynauld souligne quant à elle l'importance du texte que proposait la firme Pathé pour les projections de la vue animée *La Vie et la Passion de Notre Seigneur Jésus-Christ*. La distinction qu'elle opère entre le « boniment » et le texte de conférence (« lecture » en anglais) repose du reste sur un critère identique à celui que l'on trouve dans les recommandations de Coissac pour les projections de vues fixes : « Alors que le boniment est donné *lumières éteintes* et superpose à la projection un discours simultané, la « conférence » est faite *lumières allumées* et a une fonction didactique très nette. »<sup>115</sup> Nous ne reviendrons pas sur la distinction terminologique proprement dite, mais l'expression mise en italique par l'historienne revêt une grande importance : l'obscurité de la salle contribue à instituer la surface lumineuse de l'écran en tant qu'objet privilégié du spectacle, et ce d'autant plus lorsque la durée octroyée à chaque vue ne dépend pas du locuteur. Inversement, l'exhibition du conférencier met l'image et la parole sur un même plan, ce qui est à l'avantage de la performance dont le caractère vivant s'impose face à la représentation

113. *Ibid.*, p. 179.

114. Germain Lacasse, « De Passions en passions : le cinéma des débuts au Québec », in R. Cosandey, A. Gaudreault et T. Gunning (dir.), *Une invention du diable? op. cit.*, p. 82.

115. Isabelle Raynauld, « Les scénarios de la Passion selon Pathé (1902-1914) », in *ibid.*, p. 136.

filmique muette. Ces circonstances de la projection ont, comme le suggère Raynauld, une incidence certaine sur la nature des liens qui s'instaurent entre le boniment et la représentation visuelle. Rick Altman repère dans les indications (également consacrées à une Passion) d'un catalogue de la firme Lubin la même différenciation entre un commentaire qui précède la visualisation et un accompagnement verbal simultané: sur le modèle de ce que propose Coissac, le premier pose le contexte en faisant montre d'une certaine érudition, alors que le second livre les informations utiles à la compréhension du récit biblique tel qu'il se déroule à l'instant même sur l'écran<sup>116</sup>. La simultanéité du boniment et du défilement des images s'apparente d'ailleurs plus à la pratique des sons «fixés» du cinéma parlant, puisque la coprésence d'un flux verbal et d'un défilement photogrammatique engage inévitablement des phénomènes de synchronisation. Il en va certes également ainsi de la projection d'images fixes, mais dans une moindre mesure: s'il est nécessaire de prévoir le moment du changement de plaques, cette opération appartient totalement à la performance que réalisent conférencier et opérateur, alors que, dans le cas des scénarios évoqué par Raynauld, les producteurs entendent augmenter leur contrôle sur des films auxquels pourraient s'attaquer les censeurs en privilégiant la forme écrite. Par ce biais s'exerce une domestication des phénomènes aléatoires et labiles qui caractérisent l'oralité du «cinéma» *parlé*.

Il se peut néanmoins que le postulat d'une subordination de l'image projetée à la parole doive être nuancé en fonction de l'objet de la présentation ainsi que de la pratique pour laquelle opte le conférencier. Le commentaire de vues documentaires pouvait fort bien partir de l'image pour développer une explication, au lieu d'illustrer celle-ci avec celle-là. Dans le quatrième chapitre de son ouvrage récent sur les «sons du muet», Rick Altman compare deux types de conférence illustrée en opposant les célèbres organisateurs de spectacle lanterniste de vues de voyage que furent John Stoddard et Burton Holmes. Sur la base d'un examen du matériel et des textes utilisés par ces conférenciers, Altman constate que Stoddard met l'accent sur la transmission d'un savoir et subordonne totalement les images (qu'il n'a pas réalisées lui-même) à son texte, alors que Holmes, également photographe, construit grâce à la mise en

116. R. Altman, *Silent Film Sound*, *op. cit.*, p. 138.



séquence des vues un espace-temps cohérent qui tend à conférer une certaine autosuffisance aux images. Dans son cas, l'adjonction du verbal à la projection permet une véritable interaction audiovisuelle qui vise à intégrer le public dans la diégèse, ainsi qu'en témoignent certaines caractéristiques linguistiques du texte (l'usage du pronom « nous » et du présent). Rick Altman surestime à mon sens passablement la façon dont une telle pratique de la continuité anticiperait le travail de montage au cinéma. Il est néanmoins vrai qu'à certains égards, les rapports verbo-iconiques instaurés par le conférencier-lanterniste ont contribué à familiariser le public avec certains modes d'accompagnement verbal simultané aux projections. Le fait que la continuité de Holmes succède dans le temps à la stricte illustration de Stoddard ne doit pas être compris comme une évolution linéaire de ce type de conférences : plutôt que de parler de « vagues successives de conférenciers »<sup>117</sup> pour évoquer les différents paradigmes régissant leur conception du spectacle, il me paraît plus adéquat de considérer que ces pratiques ont été mises en œuvre de façon *parallèle*, de même que la « série culturelle » dans laquelle elles s'inscrivent entretient des liens avec d'autres séries concomitantes.

Contrairement au bonimenteur de vues animées qui visait à offrir un supplément sémantique destiné à compenser la fugacité des informations iconiques, le conférencier des spectacles de lanterne pouvait théoriquement se passer d'indications portant sur les détails de l'image, le public ayant le temps de les découvrir par lui-même. En ce qui concerne les plaques confectionnées sans recours à la technique photographique – seules images disponibles avant la généralisation du Stéréopticon dans la seconde moitié du XIX<sup>e</sup> siècle, et qui continuèrent d'être en usage parallèlement à ce dernier<sup>118</sup> – cette question de la sélection d'informations par le verbal a également trait à la nature de l'image : la peinture induit une schématisation qui, dans une certaine mesure, permettait à l'explicateur de ne pas mettre en évidence tous les éléments visuels

117. *Ibid.*, p. 55.

118. Selon Ch. Musser, *The Emergence of Cinema, op. cit.*, pp. 29-38. Musser note que le passage à la photographie introduisit d'importants changements dans les pratiques de projection : alors que les images avaient été associées jusque-là aux fantasmagories, elles offraient désormais une base plus scientifique (p. 32). Il est fort probable que l'utilisation du Stéréopticon influença également les discours qui accompagnaient les projections : les représentations détaillées et la valeur indiciaire des photographies reproduites sur les plaques de verre ouvraient la voie à la transmission de connaissances sur le monde et à l'assignation généralisée d'une certaine fonction à la parole (commentaire) et à l'image (illustration).

pertinents au niveau narratif. Par contre, l'image (ciné)photographique peut donner lieu à une richesse informationnelle quasi infinie, et appelle cette fonction du verbal que Roland Barthes a qualifiée d'*ancrage*<sup>119</sup>.

À l'époque du cinéma des premiers temps, le film était soumis à des conditions de présentation qui comportaient notamment la voix. Le verbal apparaissait donc également avant et après la projection, ce qui diminuait fortement le rôle de détermination temporelle du «film» à l'échelle du spectacle dans son intégralité. Cette relativisation des contraintes imposées par la durée filmique nous ramène à l'hétérogénéité du spectacle. Les projections cinématographiques empruntent aux spectacles scéniques qui les ont précédées (et dans lesquels elles se sont elles-mêmes inscrites) une certaine façon de présenter oralement un spectacle. Dans les Passions, on constate par exemple que le bonimenteur cinématographique renoue avec le récitant des mystères des XV<sup>e</sup> et XVI<sup>e</sup> siècles. C'est pourquoi son discours devait également présenter une dimension narrative: Burch a souligné combien ce «genre» avait contribué à linéariser la succession des vues, tout en renforçant paradoxalement, en raison de l'intertexte qu'il convoque, les caractéristiques du «tableau primitif» isolé<sup>120</sup>. Gardies dit par exemple d'une vue intitulée *La Cène* qu'elle «apparaît comme la mise en images d'un texte narratif effectivement absent, mais présent mémoriellement», et que «chaque geste ou menu événement renvoie à tel ou tel passage du récit sacré qui, en retour, fait de ce film non pas un simple geste monstatif mais une véritable narration»<sup>121</sup>. Même si cette narrativité insufflée de l'extérieur dans le film relève, dans les termes de Gaudreault, de la seule *monstration* et non de la narration, il paraît évident que le bonimenteur a bien souvent dû se substituer à l'activité mémorielle du spectateur pour actualiser le récit évangélique (ou ses multiples ramifications hypertextuelles) virtuellement présent dans les actions de l'écran<sup>122</sup>. Burch

119. Roland Barthes, «Rhétorique de l'image», *Communications*, N° 4, 1964.

120. Noël Burch, *La Lucarne de l'infini. Naissance du langage cinématographique*, Paris: Nathan, 1991, p. 140.

121. André Gardies, «Quand la narrativité vient aux films: les vues Lumière», in Jean Arrouye, Marie-Claude Taranger (dir.), *Rencontres. Croisements. Emprunts. Méthodologies de l'analyse d'images*, Aix-en-Provence: Publications de l'Université de Provence, 1996, p. 152.

122. Cet intertexte évangélique n'intervient pas seulement dans le cas de Passions filmées, mais plus généralement pour toutes les bandes qui s'inspirent des Évangiles. Sur ces questions, voir les travaux réalisés dans le cadre du projet «Les usages de Jésus au XX<sup>e</sup> siècle» (soutenu par le Fonds national de la Recherche suisse) que je codirige avec Philippe Kaenel et Jean Kaempfer à l'Université de Lausanne (<http://www.unil.ch/usagesdejesus>).

émettait déjà l'hypothèse selon laquelle de telles vues étaient « commentées par un conférencier venu tout logiquement du spectacle d'origine »<sup>123</sup>, c'est-à-dire de la forme théâtrale des mystères médiévaux. On ne peut évidemment pas postuler une filiation directe entre boniment et mystères puisque plus de quatre siècles séparent ces deux pratiques, qui sont en outre inscrites dans un contexte socioculturel totalement différent. Ce rapprochement révèle néanmoins combien les débuts du cinématographe marquèrent la résurgence d'une oralité dont certaines mutations technologiques, économiques et esthétiques viendront par la suite fortement amoindrir l'importance. En effet, la généralisation des sons fixés simultanés à la projection, la mainmise croissante de l'instance de production sur le travail de l'exploitant ou la constitution d'une narration filmique autonome basée sur le montage seront autant de facteurs qui assoiront l'hégémonie de l'*écrit*.

En ce qui concerne la discontinuité du signifiant visuel et la non-complétude du récit en images, le boniment a constitué un facteur de linéarisation, tant pour la lanterne que pour les projections animées. Le passage d'une « scène » à l'autre nécessitait, dans les premiers films « pluriponctuels » (c'est-à-dire composés de plusieurs « plans »), l'éclairage d'un commentateur averti qui maîtrisait les inférences intertextuelles ou encyclopédiques, ou du moins qui avait déjà visionné l'ensemble du métrage ; la projection fixe, fondamentalement plus discontinue, exigeait la même « béquille » orale dès lors qu'il s'agissait de construire un récit. Même lorsqu'on augmentait le nombre d'images projetées, comme le fit Alexander Black en 1894 pour son « Picture Play » intitulé *Miss Jerry* dans lequel les vues fixes se succédaient en fondus enchaînés au rythme de quatre par minute, le spectacle visuel était accompagné de musique et d'un « narrateur » (Black lui-même)<sup>124</sup>. L'ambition d'Alexander Black de raconter en images un récit complexe ne le conduisit pas à écarter la parole : l'usage de celle-ci à des fins de narrativisation est une évidence depuis sa généralisation dans les spectacles de lanterne magique où, comme le suppose Laurent Mannoni, « les ellipses trop importantes obligent certainement le lanterniste à expliquer oralement l'histoire, de manière plus détaillée »<sup>125</sup>.

123. N. Burch, *La Lucarne...*, *op. cit.*, p. 139.

124. Selon H. Geduld, *op. cit.*, p. 26.

125. L. Mannoni, *Le Grand Art...*, *op. cit.*, p. 277.

Le verbal écrit figurant sur certaines plaques ne suffisait donc pas à clarifier le récit visuel. Il est cependant intéressant que celui-ci ait été utilisé concurremment à la parole orale. Pour sa chambre obscure, Athanase Kircher avait déjà conçu un «alphabet catoptrique», le texte écrit à l'envers parvenant au spectateur après une suite de réflexions dans des miroirs. Une illustration tirée de *Ars magna lucis et ombrae* (1646), reproduite dans l'ouvrage de Laurent Mannoni<sup>126</sup>, témoigne clairement de la volonté de projeter simultanément verbal écrit et représentation iconique. On retrouvera cette double exigence du *voir* et du *lire* dans le sous-titrage des films parlants, plus proche de la pratique de Kircher que les intertitres du muet, car ces derniers se substituent généralement à l'image photographique, à moins qu'ils ne se détachent sur un arrière-fond iconique, souvent simplement dessiné (et donc non sans parenté avec les plaques de verre). L'insertion du verbal connut également d'autres modes. Mannoni dit par exemple du travail d'Auguste Lapiere, confectonneur de plaques vers 1850, qu'«entre les images sont intercalés des textes encadrés qui évoquent «l'intertitre» explicatif du cinéma muet»<sup>127</sup>. Par ailleurs, l'insertion du verbal dans les plaques croise, en France du moins, une autre «série culturelle» qui s'autonomisera à terme dans la «bande dessinée»: il s'agit des images d'Épinal, dont certains aspects graphiques témoignent, selon Perriault, de l'influence des spectacles de lanterne<sup>128</sup>. Aux États-Unis, ce sont les *comics*, dont la diffusion massive dans les journaux à grand tirage est contemporaine de l'apparition des projections cinématographiques, qui instituent un mode dominant d'utilisation conjointe du verbe et de l'image. Bien qu'elle ne fût jamais commercialisée, l'invention de Charles F. Pidgin destinée à «faire parler» les films muets est révélatrice de ce contexte d'influences intermédiateuses: il envisageait d'introduire des phylactères dans l'image lorsque les acteurs étaient censés s'exprimer<sup>129</sup>.

La comparaison avec ces autres séries culturelles n'est toutefois plus valable si l'on considère la lanterne non pas seulement comme un jouet optique, mais comme un spectacle. En effet, son dispositif se distingue radicalement de la consommation individuelle de planches comportant une série d'images: la médiation humaine

126. *Ibid.*, p. 34.

127. *Ibid.*, p. 273.

128. J. Perriault, *op. cit.*, p. 66.

129. R. Altman, *Silent Film Sound, op. cit.*, p. 168.

fait intervenir une dimension orale qui change le statut de l'écrit. Cette dimension collective et interactive est évidente dans le cas des *song slides*: les paroles qui apparaissent à l'écran renvoient à ce que chante le bonimenteur, et sont destinées à être reprises par le public. Mais l'ajout d'une parole chantée lors de ce type de projection est plus ancien, puisque les «Pantomimes lumineuses» conçues par Émile Reynaud pour son Théâtre optique et présentées au Musée Grévin dès octobre 1892 étaient déjà accompagnées, dans le cas de la série d'images intitulée *Pauvre Pierrot*, d'une musique et même, d'après ce qu'indique le scénario conservé, d'une sérénade chantée<sup>130</sup>. La chanson n'y avait toutefois pas l'autonomie qu'elle possède dans les *song slides*, puisqu'elle était inscrite dans un contexte narratif: Pierrot n'interpelle pas le public du spectacle de Reynaud, mais adresse sa sérénade à un personnage diégétique, Colombine. L'exemple de l'œuvre d'Émile Reynaud, artiste engagé par ce temple du divertissement qu'était le Musée Grévin, montre qu'en dépit des intentions généralement didactiques des «conférenciers» de lanterne magique ce mode de présentation n'était en lui-même pas dépourvu d'une dimension ludique et spectaculaire. Bien qu'il prône une utilisation «sérieuse» des projections, Coissac conseille d'ailleurs, comme on l'a vu, de ponctuer les séances par des entractes. Or, il les envisage précisément comme suit:

«C'est alors qu'on pourra faire intervenir des chants ou des déclamations, depuis le cantique jusqu'à la chansonnette, depuis le morceau de poésie jusqu'au monologue comique, suivant les circonstances.»<sup>131</sup>

En marge des projections, la voix pouvait donc, au sein de séances organisées à des fins d'enseignement, se manifester également sous des formes à dominante «attractionnelle». De plus, la lanterne a connu des utilisations diverses, puisqu'on pouvait aussi bien la trouver dans les mains d'un forain itinérant, d'un promoteur de spectacles de vaudeville, d'un professeur, d'un prêtre, etc.

Le contexte forain revêt une importance particulière puisqu'il fut également le berceau de l'exploitation du cinématographe.

130. Voir Jacques Deslandes, *Histoire comparée du cinéma*, tome I, Paris: Casterman, 1966, p. 292.

131. G.-M. Coissac, *op. cit.*, p. 201.

Or il s'agit là d'un univers bruyant où la parole est omniprésente, y compris en dehors des baraques où le « bonisseur » – parfois remplacé par un piano mécanique – harangue la foule, comme ces colporteurs de lanterne magique qui, selon Mannoni, « errent dans les rues en criant leur programme »<sup>132</sup>. Cette activité de promotion orale de l'objet visuel et ses multiples prolongements lors de la projection correspondent au degré maximal de ce que je propose d'appeler la « voix-attraction ». Théodore de Banville imite, en guise d'introduction à son recueil *La Lanterne magique* (dénommé ainsi afin d'associer les courts textes de l'ouvrage, qualifiés de « tableaux rapides », à des vues fixes), ce ton oral caractéristique des montreurs ambulants. On y trouve l'air chantonné (le prologue s'ouvre sur « tralala, deri, traderi, dère... »), la gouaille emphatique de l'annonce (« Vous verrez... ») et la vivacité de l'adresse (« Demandez la Curiosité! »)<sup>133</sup>. Ainsi les boîtes à images, utilisées dans un tel contexte, ne peuvent être rapportées à la seule dimension visuelle induite par leur machinerie. Les interventions du bonisseur devaient permettre un grand nombre de configurations dans le dispositif mis en œuvre. Mannoni constate qu'à partir de 1789, les textes des colporteurs sont plus fortement ancrés dans la situation politique et prennent un tour pamphlétaire<sup>134</sup>. On retrouve là ce potentiel de « résistance » que Germain Lacasse voit dans les faits d'oralité en général, et dans le boniment de cinéma en particulier.

Parmi les diverses utilisations de la lanterne magique, l'une d'elles accorde à la voix un statut particulier que l'on retrouve dans le cinéma parlant, par exemple lorsque certaines voix détachées des corps dont elles sont censées émaner nous transportent dans le domaine de l'occulte. Ce mode de projection pré-cinématographique est celui de la fantasmagorie, un type de spectacle qui s'est répandu à la fin du XVIII<sup>e</sup> siècle et dont le principal promoteur fut Robertson. Comme d'autres dispositifs optiques avant elle, la fantasmagorie pouvait être produite et perçue sur le mode de la crédulité, ou dans une optique éducative à la Kircher. Jann Matlock souligne la persistance de cette ambiguïté même lorsque les intentions de l'organisateur du spectacle étaient claires : « La « Fantasmagorie » [...] se présentait explicitement comme une tentative de dévoiler

132. L. Mannoni, *Le Grand Art...*, op. cit., p. 81.

133. Théodore de Banville, *La Lanterne magique*, Paris: Charpentier, 1883, p. 3.

134. L. Mannoni, *Le Grand Art...*, op. cit., p. 99.

les duperies des prêtres et des illusionnistes (même si son déroulement avait pour effet ultime de convaincre nombre de ses spectateurs que les fantômes existaient vraiment)... »<sup>135</sup> La question de l'illusion (affichée ou niée) contribue dans une large mesure à définir cette pratique spectaculaire. Or l'étymologie du terme « fantasmagorie » (ou « phantasmagorie ») est révélatrice d'une association fondamentale entre cette fonction illusionniste et la parole : il s'agit de faire illusion (« phantazô ») et d'engager un dialogue avec des spectres (« agoreuô », je parle)<sup>136</sup>. La voix constitue donc un élément définitoire de ce spectacle que les historiens du pré-cinéma qualifient traditionnellement d'« optique ». Elle guide les spectateurs à travers l'univers macabre dans lequel ils sont immergés grâce à un dispositif qui produit des effets d'étrangeté (création de fumée, sons lugubres d'un « harmonica de verre », dissimulation de la source des images). Les effets de la voix sur le public des fantasmagories varient en fonction des intentions du conférencier et du régime de croyance de l'audiospectateur.

D'une part, la voix permet de renforcer l'impression de réalité du dispositif lorsqu'elle prend pour objet la *représentation* : comme les projections ont lieu dans l'obscurité, la voix du fantasmagore participe du caractère immatériel d'un spectacle qui se veut spectral. La non-visibilité de la source vocale confère à la voix un statut « flottant » qui peut être rattaché aux fantômes représentés visuellement, ce qui renforce le pouvoir illusionniste du spectacle. Si l'on se réfère à *La Magie blanche dévoilée* (1784) d'Henri Decremps, qui décrit une séance de lanterne magique telle qu'en réaliseront les fantasmagores, on comprend que la voix constituait un élément central au sein de la débauche d'effets visuels, auditifs et même kinesthésiques produits par le « magicien ». L'auteur évoque notamment « une voix de tonnerre qui fait trembler les vitres, en prononçant ces mots : « Insensé, qui ne crois qu'à la blanche Magie, tremble ! Voici l'enfer avec sa diablerie. »<sup>137</sup> Les voix sont de toute évidence utilisées en tant qu'effets sonores, et en cela ne se distinguent pas vraiment des bruits tonitruants qui retentissent dans la salle. Cependant, le langage articulé intervient également sous la forme d'une adresse, postulant un locuteur

135. Jann Matlock, « Voir aux limites du corps : fantasmagories et femmes invisibles dans les spectacles de Robertson », in Le Men (dir.), *Lanterne magique...*, op. cit., p. 86.

136. Selon L. Mannoni, *Le Grand Art...*, op. cit., p. 135.

137. Cité par L. Mannoni, *ibid.*, p. 136.

interne à la représentation, qui, en s'adressant à l'audiospectateur, le happe dans celle-ci. Or cette voix enjoint à une croyance en la nature magique de la représentation, la parole se faisant performative puisque sa présence même « atteste » qu'elle provient de l'au-delà. Cette « voix de tonnerre » n'est pas celle d'un éventuel présentateur, mais elle est partie intégrante des éléments (sur)naturels qui se déchainent, de l'environnement magique dans lequel le spectateur est immergé. Désincarnée, elle se met au service des effets produits par le dispositif. Jérôme Prieur émet d'ailleurs l'hypothèse selon laquelle certaines séances de fantasmagorie auraient donné lieu à des performances de ventriloquie<sup>138</sup>. Cette idée est symptomatique de la séparation du corps et de la voix qui s'y joue : le locuteur qui se déplace dans le noir se fait l'équivalent du projecteur dissimulé derrière l'écran pour la rétro-projection mobile. L'incarnation est d'autant plus « vraisemblable » que, à l'instar de la voix, son support visible est évanescent. Une phrase tirée des *Mémoires* de Robertson révèle combien le fantasmagore savait user des connotations surnaturelles de la voix proférée dans l'obscurité : « Il frappe au troisième tombeau ; une ombre se lève et lui demande : « Que me veux-tu ? »<sup>139</sup> Le passage de la troisième personne au discours direct indique une volonté de présentifier la narration, ce à quoi l'oralité est particulièrement apte. Il se peut bien que certains audiospectateurs aient pris la voix du commentateur pour celle d'un esprit. Notons d'ailleurs que ce régime de croyance était requis par les séances de spiritisme qui connurent précisément un essor durant la seconde moitié du XX<sup>e</sup> siècle, lorsque débute l'ère de la fixation et de la reproduction mécaniques des sons.

D'autre part, le discours du conférencier pouvait se référer à la *machinerie*. Tom Gunning a relevé que la parole agit dans les présentations de Robertson comme un facteur de démystification du spectacle en expliquant les principes scientifiques sur lesquels repose l'illusion :

« La fantasmagorie créait volontairement une zone de tension entre crédulité (plusieurs membres du public étaient convaincus que la salle abritait de véritables revenants) et la démystification,

138. Jérôme Prieur, *Séance de lanterne magique*, Paris : Gallimard, 1985, pp. 136-137.

139. Cité par L. Mannoni, *Le Grand Art...*, *op. cit.*, pp. 157-158.



le spectacle étant présenté par le conférencier comme une curiosité optique, explicable par des principes scientifiques.»<sup>140</sup>

Parce qu'il traite essentiellement de la dimension optique du dispositif, Tom Gunning place le «fantasmagore» – qu'il nomme significativement «conférencier» – du côté de l'explication, et ne fait référence à aucune instance vocale pour le pôle inverse, celui de la crédulité. Cette partition entre image et voix est peut-être à nuancer, puisque la tension qui s'établit selon lui entre crédulité et incrédulité n'est pas limitée au partage opéré entre les effets optiques et sonores d'une part, et la voix d'autre part. Cette bipolarité traverse en fait l'ensemble des possibilités d'usage de la voix, qui peuvent varier d'un spectacle à l'autre ou même au sein d'une même séance. D'ailleurs, tous les lanternistes qui s'inspiraient de Robertson n'étaient probablement pas aussi regardants en ce qui concerne la dimension éducative de cette activité.

Ces quelques remarques à propos de la pratique lanterniste permettent de relativiser le caractère singulier de l'oralité du «cinéma» parlé, les images produites par la lanterne magique ayant probablement toujours appelé la présence de la voix. En 1622, Constantin Huygens, père de Christiaan, jésuite protestant qui jouera un rôle important dans la propagation de la lanterne, faisait l'éloge d'une projection (probablement réalisée à l'aide d'une version primitive de la lanterne magique) à laquelle il avait assisté en affirmant que «c'est icy la vie même, ou quelque chose de plus relevé, *si la parole n'y manquoit*»<sup>141</sup>. La (re)production d'une image appartenant à un dispositif spectaculaire semble de tout temps avoir suscité le désir d'adjoindre la voix pour donner «vie» à la représentation. Les séances des premiers temps du cinématographe n'échappaient pas à cette règle lorsqu'elles faisaient intervenir un bonimenteur.

### 1.3 VOIX VIVE ET SYNCHRONISATION

Avant d'approfondir la question de la médiation bonimentaire, j'aimerais m'attarder sur le pôle de la représentation audiovisuelle dans le «cinéma» parlé. Le caractère aléatoire de l'accompagnement des vues animées n'exclut pas que le bonimenteur ait pu

140. Tom Gunning, «Fantasmagorie et fabrication de l'illusion: pour une culture optique du dispositif cinématographique», *Cinéma*, Vol. 14, N° 1, 2003, p. 73.

141. Cité par L. Mannoni, *Le Grand Art...*, *op. cit.*, p. 43 (je souligne).

procéder à une synchronisation, même si la précision de cette dernière ne correspondait peut-être pas toujours aux standards ultérieurs qu'imposèrent les systèmes mécaniques à travers le modèle de la «voix synchrone». C'est pourquoi il est préférable de distinguer la *synchronisation* (voco-iconique), opération résultant de l'instance de production de la représentation, du *synchronisme*, quant à lui relatif à l'appréhension subjective de l'audiospectateur. S'il est difficile de définir les normes du synchronisme qui prévalaient auprès du public des projections du début du XX<sup>e</sup> siècle, on peut par contre tenter d'évaluer sur quelles bases (et dans quelles limites) l'exercice de la voix vive participait d'une entreprise de synchronisation. L'«espace de la production» étant généralement scindé, dès les projections de lanterne magique, entre un orateur situé à proximité de l'écran et un opérateur chargé de la manipulation de l'appareil, il va de soi qu'une gestion conjointe de ces deux activités était indispensable. G.-M. Coissac souligne par exemple la nécessité de recourir à un «opérateur exact et attentif, qui suit la conférence avec soin, de façon à changer la vue en temps opportun»<sup>142</sup>. Si, dans le cas de telles projections fixes, ce sont les apparitions d'images qui se calquent sur le discours du conférencier, le déroulement temporel de la séance n'en est pas moins planifié.

Notons que la synchronisation n'a pas subi de progression uniforme vers un degré croissant de mécanicité, puisque la voix vive s'est perpétuée dans plus d'un dispositif «multimédia» après la généralisation du parlant, tandis que certaines techniques de couplage avec le phonographe ou avec des instruments de musique mécaniques ont existé depuis les tout premiers temps des projections cinématographiques. L'opposition de l'humain et de la machine concerne néanmoins en propre la synchronisation qui, idéalement, exige de la voix humaine un minutage exact que l'on associe plutôt à une certaine «mécanicité».

Une réflexion sur la synchronisation ne peut à mon sens faire l'économie d'une définition des principes sur lesquels repose la mise en relation «temporelle» de la parole et l'image. La présence, au niveau linguistique, d'unités phonétiques et de signification rend la question du découpage syntagmatique de la chaîne verbale particulièrement aisée. Il en découle en effet une délimitation d'un

142. G.-M. Coissac, *op. cit.*, p. 183.

segment dont la longueur, variable, définit le degré de précision de la synchronisation. Ces unités sont les suivantes : le phonème, le mot, le syntagme linguistique, la phrase et le texte dans son intégralité. Alors que la notion de « phonème », issue de la phonologie, est seule capable de prendre véritablement en compte la dimension orale des phénomènes, les autres unités appartiennent à l'analyse linguistique d'énoncés qui peuvent tout aussi bien être écrits. Je préfère néanmoins opter pour le terme courant de « mot » au lieu de « morphème », car il renvoie à une unité plus nettement repérable dans le flux oral. Le syntagme est quant à lui lié à la structure interne de la phrase qui, dans la majorité des cas, influe sur la perception du lecteur (ou de l'auditeur), même s'il est toujours permis au locuteur d'opérer lui-même une segmentation indépendante. Le découpage phonématique est essentiel pour l'établissement d'une synchronisation que l'on appellera *stricte* (la durée d'un phonème se mesurant en centièmes de seconde), car il permet une multiplication des repères sur un faible intervalle temporel ; à l'autre extrême, l'unité du texte impose uniquement les deux seuils qui le délimitent, sa lecture courant parallèlement au défilement des images selon une synchronisation *large* sans afficher d'autres *points de synchronisation*, ainsi que Michel Chion propose de nommer ces instants de convergence entre les pistes visuelle et auditive<sup>143</sup>. On parlera de « segment » lorsque l'unité n'est pas aussi discrète que le « point ».

Le langage parlé n'est pas exclusivement constitué d'unités porteuses de signification. Il importe également de considérer d'autres aspects dits « prosodiques », justement qualifiés de « suprasegmentaux » parce qu'ils échappent à toute segmentation entreprise à partir des catégories de l'analyse linguistique<sup>144</sup>. L'introduction de variations dans l'un des facteurs de la prosodie peut en effet créer des points de synchronisation, pour autant que ces modulations soient perceptibles par l'audiospectateur. L'accentuation, en tant que phénomène de mise en relief d'un élément discret, constitue en quelque sorte le principe général de toute ponctuation orale. La pause fait également office de ponctuation, offrant ainsi un segment synchronisé potentiel. La relative labilité des énoncés oraux permet d'affranchir quelque peu la distribution des accents par rapport aux contraintes

143. M. Chion, *L'Audio-vision, op. cit.*, p. 52.

144. Voir André Martinet, *Éléments de linguistique générale*, Paris : Armand Colin, 1970, p. 21.

imposées par la structure syntaxique. Umberto Eco note l'existence d'une latitude « idiosyncrasique » dont dispose le locuteur en dépit des règles phonologiques propres à une langue donnée, de sorte que chaque utilisateur colore les « différents phonèmes prévus par sa propre langue » en recourant à des « intonations individuelles »<sup>145</sup>. Plus les divergences par rapport aux « normes » de la langue concernée sont importantes, plus l'unité verbale est marquée, et donc susceptible de provoquer une perception sur le mode du synchronisme. D'autres caractéristiques orales peuvent jouer un rôle dans l'établissement d'un point de synchronisation dès lors qu'elles subissent un changement notable: il en va par exemple ainsi du rythme (lié au nombre et à la durée des pauses, au débit, etc.), de l'intonation, de l'intensité (souvent corrélée à la hauteur des sons) ainsi que de tous les traits de la voix elle-même (accent, grain, timbre, ...). Toute variation suprasegmentale fait office de démarcation, et peut donc constituer le lieu d'une synchronisation verbo-iconique.

À ce découpage de la parole correspond la sélection d'unités discrètes dans l'image. Sur ce point, l'opération est passablement moins réglementée puisque, comme l'a démontré Christian Metz, il n'y a pas de « double articulation du langage » au cinéma<sup>146</sup>. Les théoriciens qui, dans la vague sémiologique des années 70, se sont évertués à débusquer dans l'image animée certains équivalents des unités linguistiques, ont prouvé à leur insu l'impossibilité d'une telle entreprise<sup>147</sup>. En ce qui concerne la synchronisation, cette « équivalence » paraît toutefois concevable dans la mesure où elle requiert uniquement une inscription temporelle de l'unité « filmique » entre deux bornes aisément perceptibles. Par conséquent, cette caractéristique est indépendante de la question du sens, alors que le problème des analogies entre langues naturelles et « langage cinématographique » résidait principalement dans le fait qu'aucune unité de l'image n'est « vide de sens », et donc qu'on ne peut trouver d'équivalents « visuels » aux phonèmes.

Il est donc possible de revenir, dans une optique différente, aux notions proposées par les théoriciens qui furent partisans de la « double articulation » au cinéma, notamment lorsqu'elles ont trait

145. Umberto Eco, « Sémiologie des messages visuels », *Communications*, N° 15, 1970, p. 23.

146. Voir Christian Metz, *Langage et cinéma*, Paris: Larousse, 1971.

147. Voir la synthèse proposée par Roger Odin, *Cinéma et production de sens*, Paris: Armand Colin, 1990, pp. 78-90.

à la question du mouvement. Pour Pier Paolo Pasolini, les unités filmiques qu'il appelle « cinèmes » sont issues du monde filmé, et notamment des « actes de la réalité ». Elles se distinguent par ailleurs hiérarchiquement de l'unité supérieure du « monème », c'est-à-dire du plan<sup>148</sup>. Cette différenciation s'applique selon moi également au repérage d'unités utiles à la synchronisation, puisqu'elles ressortissent soit au représenté (les mouvements perceptibles à l'écran), soit à la représentation (la gestion temporelle de la succession des plans au montage, à laquelle il faut ajouter une dimension que Pasolini occulte en associant représenté et réalité : les mouvements de caméra au tournage). L'unité du premier type est prise en compte, dans le cas spécifique du mouvement, par la théorie contemporaine et parente édifiée par Umberto Eco, qui se réfère à Pasolini tout en faisant preuve, dans une démarche sémiologique plus orthodoxe et cohérente, d'une conscience du caractère intrinsèquement conventionnel des « actions de la réalité ». À la suite du théoricien Ray Birdwhistell, qui travaille sur la sémiologie de la gestuelle humaine (la « kinésique »), Eco identifie notamment, dans le « continuum iconique » qu'offre le médium cinéma, l'unité discrète du « *kine* », définie par la décomposition photogrammatique du geste d'un personnage<sup>149</sup>. Selon ce modèle, les *kines*, dépourvus de signification, se combinent pour donner lieu à une unité sémantique plus vaste dénommée « kinémorphe », c'est-à-dire à une unité gestuelle signifiante susceptible d'être étudiée par la science kinésique. Considérés indépendamment de la théorie de la signification quelque peu obsolète dans laquelle ils s'inscrivent, les *kines* et *kinémorphes* discutés par Umberto Eco me semblent constituer des notions utiles pour rendre compte de certains processus régissant l'établissement de *points* et de *segments* de synchronisation. Un repérage d'unités iconiques basé sur ces catégories permet en effet d'esquisser une typologie des modes de synchronisation.

Dans le cinéma parlant, l'ancrage visuel du point de synchronisation s'effectue dans la grande majorité des cas à partir des mouvements représentés à l'image, la pratique dominante résidant dans la synchronisation *vocolabiale*. Le mouvement des lèvres des individus visualisés ne constitue cependant qu'un type de mouvement

148. Pier Paolo Pasolini, « La langue écrite de la réalité » [1966], in *L'Expérience hérétique*, Paris: Payot, 1976, pp. 171-173.

149. U. Eco, « Sémiologie... », *art. cit.*, pp. 44 et 46.

parmi d'autres; il est certes nécessaire de le désigner en propre puisqu'il est la condition sine qua non du «synchronisme», mais il faut aussi élargir cette question à tous les types de *kines*. En fait, le mouvement dans l'image constitue un facteur plastique indépendant du caractère figuratif de l'image. D'ailleurs, les films sonores abstraits procèdent également à une synchronisation, et tendent même à exhiber ce principe en développant des enjeux purement rythmiques. Néanmoins, cet usage du son concerne avant tout la musique, la voix appelant, comme on l'a vu, la figuration de l'humain. La prédominance de l'anthropocentrisme nous incite par contre à isoler un autre type de mouvement que celui de la bouche: le geste. On passe alors d'un rapport de causalité (fictive) entre la source vocale et le mouvement à un facteur purement temporel: une parole est émise simultanément à la monstration d'un geste, généralement interprétable en tant que kinémorphe par le spectateur sur la base de codes extracinématographiques. Le geste peut être propre au locuteur (voix synchrone) ou effectué par d'autres personnages (il se désolidarise alors du lieu d'émission). Dans le premier cas, cette *synchronisation vocogestuelle* vient en renfort de celle qui s'opère au niveau vocalabial (à moins que la bouche du locuteur ne soit dissimulée) puisque, lorsque le locuteur n'est pas filmé en gros plan, la corrélation de la parole et des gestes contribue à souligner, voire à «authentifier» la synchronisation. Ainsi, selon les dires de W. K. L. Dickson, le premier *talkie* mythique qu'il aurait projeté en 1889 à Edison l'aurait montré lui-même s'adressant à son patron, puis comptant jusqu'à dix en levant puis en baissant la main à chaque chiffre<sup>150</sup>. Bien que probablement mensonger<sup>151</sup>, ce propos est révélateur de la façon dont l'inventeur avait imaginé prouver l'exactitude que son invention présentait en termes de synchronisation.

L'audiospectateur remarque d'autant plus la synchronisation vocogestuelle que les composantes visuelles et auditives sont exhibées dans leur dimension ponctuelle, «métrique»: ainsi, l'association de la parole chantée et des gestes chorégraphiés représente-t-elle la forme la plus aboutie de ce type de synchronisation. Non seulement la mélodie permet une juxtaposition de rythme, mais les paroles elles-mêmes sont souvent soumises à des conventions

150. Cité par H. Geduld, *op. cit.*, p. 17.

151. Voir R. Altman, *Silent Film Sound*, *op. cit.*, pp. 78-79.

poétiques (rimes, répétitions, césures, etc.) qui influent sur la structure prosodique du texte. Lorsque la phrase se fait phrasé, c'est-à-dire un « équivalent dans la musique de ce qu'est [...] le geste dans l'action »<sup>152</sup>, les concordances les plus subtiles sont possibles. En outre, les mouvements « profilmiques » ne se réduisent pas aux gestes des personnages : tout objet qui se meut dans l'image peut donner lieu à des points de synchronisation visuelle. Je nommerai cette catégorie plus large *synchronisation (voco)cinématique*. Un mot peut être prononcé au moment même où un véhicule s'arrête, où un objet tombe, où une machine se met en mouvement, etc. Ce genre de synchronisation concerne néanmoins plus souvent les bruits que les voix.

À ces mouvements *dans* l'image, il faut ajouter les mouvements *de* l'image, qui peuvent également occasionner une synchronisation voco-iconique. Ces derniers relèvent de deux types : certains sont issus du déplacement d'éléments profilmiques (ou donnés comme tels, comme dans les films d'animation ou dans les images de synthèse), d'autres sont liés à la mise en cadre. Les seconds sont faciles à concevoir : un travelling, un panoramique ou un zoom présentent nécessairement des seuils, plus ou moins perceptibles en fonction de leur vitesse, du caractère plus ou moins progressif ou brusque de leur amorce, de la nature des mouvements profilmiques susceptibles d'interagir avec eux, etc. L'instant le plus marqué est celui du passage de l'immobilité au mouvement du cadre. À ces bornes externes peuvent s'ajouter nombre de seuils internes définis par des changements de vitesse, de direction, etc. En tenant compte de la terminologie introduite par Eisenstein et popularisée par André Gaudreault, je propose de nommer ce procédé *synchronisation (vocale) de mise en cadre*.

Le dernier niveau, qui est convoqué par ce que j'appellerai, en recourant au même cadre terminologique, la *synchronisation vocale de mise en chaîne*, comprend tous les types qui mobilisent l'articulation même des images. L'unité visuelle est essentiellement le plan – le photogramme n'étant pas une unité pertinente puisque sa durée de  $\frac{1}{4}^e$  de seconde permet l'émission d'à peine plus de quelques phonèmes – le point de synchronisation correspondant à ses seuils. L'un d'eux peut en outre marquer les bornes d'une unité plus grande, par exemple lorsqu'il s'agit du premier ou du

152. André Wyss, *Éloge du phrasé*, Paris : PUF, 1999, p. 281.

dernier plan d'une séquence. En adaptant ses propres segmentations à celles des images, la parole souligne, dans le cas de la synchronisation de mise en chaîne, le caractère fragmentaire du signifiant visuel. Toutefois, les *cuts* du montage n'étant pas systématiquement affichés en raison de la présence simultanée d'une unité linguistique, la parole contribue plutôt, globalement, à renforcer la continuité fictive du flux visuel. Les points de synchronisation visuels sont manifestes dans des plans qui font office, selon la terminologie metzienne adéquate dans ce contexte, de «signes de ponctuation», comme les fermetures ou ouvertures à l'iris, les fondus, les images noires ou blanches, les cartons de titre, etc. qui soulignent le caractère discontinu du film<sup>153</sup>. La synchronisation de mise en chaîne induit donc une réflexivité plus importante que les autres modes de synchronisation qui, définis quant à eux en fonction du représenté, n'attirent pas l'attention sur les qualités physiques du support. La synchronisation de mise en chaîne est le seul type qui peut être utilisé lorsque la représentation visuelle est composée d'images fixes, comme cela est le cas dans le film *La Jetée* de Chris Marker (1962).

On pourrait dire qu'à l'époque du «cinéma» parlé, la synchronisation vocogestuelle jouait un rôle similaire à celui de l'accompagnement musical. François Jost remarque, à propos de la musique écrite par le compositeur Camille Saint-Saëns pour *L'Assassinat du duc de Guise*, «que deux missions qui furent celles du conférencier incombent maintenant à la musique, celle de ponctuer les mouvements et les gestes, et celle de constituer des unités syntagmatiques et de les distinguer les unes des autres»<sup>154</sup>. S'il est prématuré de fixer l'obsolescence du boniment en 1908 – les recherches récentes ont montré qu'il a eu cours en France jusque vers le milieu des années 10<sup>155</sup> et au Québec jusque dans les années 20<sup>156</sup> – et, s'il faut tenir compte du caractère exceptionnel de cette composition originale ainsi que des conséquences de l'absence de «fixation» de cet accompagnement, il est vrai que musique et voix, pro-

153. Christian Metz, *Essais sur la signification au cinéma*, tome 2, Paris: Klincksieck, 1972, pp. 111-137.

154. François Jost, «The Voices of Silence», in R. Abel, R. Altman (dir.), *The Sounds of Early Cinema*, *op. cit.*, p. 54.

155. F. Albera et A. Gaudreault, «Apparition, disparition et escamotage...», *art. cit.*, p. 184.

156. André Gaudreault et Germain Lacasse, «L'écran ventriloque», in A. Gaudreault (dir.), *Au pays des ennemis du cinéma... Pour une nouvelle histoire des débuts du cinéma au Québec*, Montréal: Nuit Blanche, 1996, p. 139.



duites simultanément à une projection d'images dans une visée de synchronisation, partagent les propriétés décrites par Jost, que l'on peut qualifier de « rythmiques ». Elles ne résultent pas uniquement de coïncidences entre la parole, la musique et les gestes visualisés, mais concerne aussi le montage : comme le montre l'usage dominant de la musique de film depuis la généralisation du parlant, les sons extradiégétiques se développent généralement en adéquation avec l'organisation des plans du film. C'est pourquoi la typologie des modes de synchronisation que j'ai proposée ci-dessus permet d'aborder à divers niveaux la question du « rythme » au cinéma, au sens restreint d'un ensemble d'interactions audiovisuelles. Il y a un rythme du montage comme il y a un rythme des actions de l'écran, celui-ci résultant d'une série de mouvements qui, eux-mêmes, procèdent d'une coordination mutuelle de la vitesse. Au niveau de la bande-son, de semblables régularités sont décelables, non seulement pour la musique (un cas si évident qu'il est souvent érigé en interprétant principal du rythme cinématographique à travers la notion de « musicalité »)<sup>157</sup>, mais aussi pour la voix.

Dans le cinéma parlant, cette fonction rythmique de la voix s'accomplit pleinement dans l'utilisation d'une voix-over, ce procédé n'étant pas asservi à la synchronisation vocolabiale. Souvent considérée à travers l'influence qu'elle exerce sur la temporalité diégétique, la voix-over demande également à être examinée du point de vue de son inscription dans le défilement du film proprement dit : en jouant de pauses, d'attaques, de variations dans le débit, d'alternances des locuteurs, etc., elle représente un facteur-clé de la perception du rythme filmique.

Ces divers paramètres valent également pour la voix vive, quand bien même la synchronisation bonimentorielle n'est pas aussi stricte que celle d'une voix qui, fixée sous la forme d'une piste optique, entretient un rapport de contiguïté physique avec la bande-image. Sujette à d'importantes variations d'une projection à l'autre, la synchronisation humaine tend certes à privilégier la composante sémantique – phénomènes relevant de la *coréférence* (verbo-iconique) – sur la question du réglage temporel : le texte évoque ce que l'image montre plutôt qu'il n'est minuté par rapport au défilement du film. Il n'en demeure pas moins que ce réglage

157. Voir notamment Jean Mitry, *Esthétique et psychologie du cinéma*, Vol. 2 (« Les Formes »), Paris : Éditions Universitaires, 1965, pp. 141-176.

des relations verticales interne à la représentation doit être pris en compte dans l'analyse de l'insertion du bonimenteur au sein du dispositif cinématographique.

À l'époque du «cinéma» parlé, la synchronisation représentait déjà l'un des enjeux de l'emprise de l'instance de production des films sur les conditions labiles de leur exploitation dans les salles. Un catalogue de la Star Film prescrit par exemple une lecture synchronisée d'un texte livré avec la bande *La Légende de Rip van Vinckle* (1905), qualifié de «résumé destiné à être lu phrase par phrase au moment exact de la projection des parties correspondantes de la vue cinématographique»<sup>158</sup>. Cette précision est d'importance, puisque la firme de Méliès prône cette «exactitude» dans la «correspondance» que cherchent à atteindre, à la même époque (avec des systèmes comme le Phonorama, le Graphophone, etc.), les nombreux inventeurs de techniques de synchronisation mécanique du son et de l'image. Toutefois, la fonction de «résumé» du texte lu implique non pas que ce dernier soit composé de dialogues devant coller aux mouvements des lèvres des acteurs, mais qu'il relève, à l'instar de celui que véhicule une voix-over, d'une instance d'énonciation ostensiblement distincte des personnages visualisés. Le réglage du défilement temporel du verbal est néanmoins pensé par rapport à l'image sur le mode de la synchronisation de mise en chaîne. En effet, une découpe en unités linguistiques («phrase par phrase») trouve un équivalent dans la succession des «tableaux» du film, qui en compte dix-sept. Le «synchronisme» est donc *large*, tant au niveau de l'unité verbale considérée qu'en ce qui concerne la latitude importante laissée au bonimenteur, puisque rien n'est précisé quant à la nature de la correspondance verbo-iconique, l'expression «parties correspondantes» pouvant également servir à désigner des actions distinctes se succédant au sein d'un même «tableau» (synchronisation vocogestuelle). Le cas le plus précis serait celui de phrases uniques énoncées à l'apparition d'un nouveau tableau, c'est-à-dire là où se présente sur la bande-image un seuil permettant la délimitation d'un segment synchronisé. Cependant, cette fonction du bonimenteur qui consiste à titrer les parties ou, à l'aide de périphrases, à fournir une synthèse des actions qui seront montrées par la suite,

158. Cité par André Gaudreault, «Le retour du [bonimenteur] refoulé... (ou serait-ce le bonisseur-conférencier, le commentateur, le conférencier, le présentateur ou le «speaker»?)», *Iris*, N° 22, 1996, p. 25.

ne constitue qu'une modalité parmi d'autres d'inscription temporelle<sup>159</sup>. En effet, le boniment peut tout aussi bien être proféré de façon continue et simultanément aux images.

La synchronisation bonimentorielle participe d'un effet pragmatique plus large que l'on peut selon moi rapprocher de la fonction de « polissage » que Noël Burch, dans une perspective différente, propose de voir dans l'activité du bonimenteur<sup>160</sup>. Le synchronisme vocal, même très large et (trop) humain, tend à donner prise à une réappropriation du verbal par la diégèse du film qui diminue l'hétérogénéité du spectacle et, partant, ramène l'attention du spectateur à l'image. Le synchronisme implique une certaine isomorphie avec l'écran qui lui permet d'asseoir un processus de linéarisation du sens. Plus la synchronisation est précise, plus le bonimenteur migre, au sein de « l'espace de la production », du pôle spectatoriel vers celui de la machinerie. En ajustant sa parole au défilement du film, il tend à se « mécaniser ». Il n'est donc pas surprenant que certains exploitants aient eu recours à des systèmes semi-mécaniques. On ne peut dès lors plus vraiment parler de « médiation » bonimentorielle dans la mesure où le texte, bien qu'énoncé oralement, s'inscrit dans une dynamique réglée à l'avance (c'est la fixation propre à l'*écrit*) et, en se mettant totalement au service du représenté, subit une intégration contraire à cette extranéité qui confère une position d'intermédiaire au bonimenteur. Les qualités de la performance vocale subsistent, mais elles sont fortement contraintes par ce principe d'unification des composantes visuelle et verbale qui culminera dans le cinéma parlant avec l'élaboration d'un « discours audiovisuel ».

Les problèmes soulevés par la synchronisation du boniment et de l'image sont discutés dans un article de G.-M. Coissac publié en 1911. Coissac avise son lecteur des risques de parasitage provoqués par une énonciation verbale simultanée aux images :

« Il semble que ces vues hachent les syllabes des mots que l'on désirerait leur appliquer. Chaque tremblement devient une virgule, un trait de sursaut qui, par les yeux, reçoit le sens de l'ouïe. »<sup>161</sup>

159. On verra toutefois au chapitre III l'importance de cette fonction de titrage.

160. Noël Burch, « Discipline et jouissance », in Bernard Chardère, Philippe Dujardin, Lucette Lombard-Valentino (dir.), *Lumière, le cinéma*, Lyon : Institut Lumière, 1992.

161. « La Conférence cinématographique », *Ciné-Journal*, N° 161, 23 septembre 1911, cité par F. Albera et A. Gaudreault, « Apparition, disparition et escamotage... », *art. cit.*, p. 188.

Pour Coissac, qui postule une sorte de transposition sensorielle du rythme visuel en stimuli auditifs, le problème de la simultanéité du verbe et de l'image réside dans l'incompatibilité qu'il perçoit entre le caractère continu du flux de parole et la discontinuité du signifiant cinématographique, ce dernier imposant son rythme saccadé au texte et entraînant la voix humaine dans la mécanicité du défilement cadencé de la pellicule. La segmentation sur laquelle repose le fonctionnement d'appareils destinés à assister la synchronisation *live* – le principe du Ciné-pupitre consiste par exemple à faire défiler lors de la projection les mots lus à voix haute par des « déclamateurs » sur une bande dont l'avancée est synchronisée avec celle de la pellicule du film<sup>162</sup> – s'opère donc selon Coissac au niveau même des processus perceptifs qu'engage la vision d'un film. C'est d'ailleurs sur la perceptibilité du « rythme » visuel par l'audiospectateur que se fonde la synchronisation vocale de mise en chaîne. Coissac va cependant plus loin en affirmant une supériorité du rythme visuel qui contribuerait à « créer » de nouvelles caractéristiques prosodiques, écrasant le phrasé de l'énoncé oral.

La prééminence de la machinerie sur la performance humaine dont parle Coissac ne peut toutefois pas être universalisée sans tenir compte de la nature de ladite performance, si ce n'est en sous-estimant largement les implications sémantiques et pragmatiques de la vocalité. Un exemple de boniment particulièrement travaillé est imaginé, probablement sur la base de souvenirs, par Blaise Cendrars dans les *Confessions de Dan Yack*. Dans ce texte

162. Voir par exemple Charles Delacommune, «Le Synchronisme de la parole, de la musique et des bruits au cinéma», in Maurice Noverre, *Septumia. Légende-mime (Mœurs antiques)*, Bordeaux; Paris: Bourlange et Mollat; Dion, 1923. On mentionne généralement l'appareil de Delacommune exclusivement pour évoquer la possibilité d'une synchronisation des images animées avec la musique. Or cette éviction des manifestations vocales, qui ne correspond pas aux intentions de l'inventeur, résulte à mon sens du contexte de réception du début des années 20, où la parole connaît momentanément une perte d'intérêt qui est en grande partie liée à l'émergence d'un discours de légitimation artistique du médium visant à déprécier tout élément considéré comme une « béquille » du discours (ou du rythme) spécifiquement créé par les images. La différence entre les explications de Charles Delacommune, qui envisage le doublage *live* pour des réalisations résolument « artistiques », et la réception que l'on fait de son système (voir par exemple Didier Monclair, «Une invention française. Le Ciné-pupitre Delacommune», *Cinémagazine*, N° 11, 16 mars 1923, p. 453) est symptomatique de l'état d'esprit régnant parmi les intellectuels de l'époque, qui privilégient les développements de la musique d'accompagnement. Le texte du scientifique témoigne au contraire d'un désir de contribuer à un progrès technique en introduisant *avant tout* la parole au cinéma, ainsi que certains en auraient manifesté le désir: «Pour quelques-uns l'idéal serait l'enregistrement et la reproduction simultanée des images et des sons, et en particulier de la parole.» En affirmant que «la parole demeure – et demeurera longtemps encore – le seul véhicule de la pensée» (*op. cit.*, pp. 15-16), Delacommune soustrait son discours au paradigme du «tout-visuel» qui prévaut chez les premiers partisans du «7<sup>e</sup> Art».

publié en cette année-clé de 1929 qui a vu s'amorcer la généralisation du parlant, le personnage éponyme lit à haute voix devant un dictaphone les cahiers de sa défunte amie Mireille. Celle-ci évoque les séances de cinéma auxquelles elle assista avec lui à partir de 1918, où un bonimenteur qui se faisait appeler « Monsieur Nouvel-An » s'adonnait à de pures performances orales :

« Pour les films de Charlot, il improvisait des sketches vertigineux et irrésistibles. Sa voix imitait toutes les contorsions, toutes les acrobaties de Charlot et de ses comparses, et dans les moments pathétiques, quand l'action marquait un temps d'arrêt [...], M. Nouvel-An trouvait des mots d'un drôle qui vous tiraient les larmes. Le petit chapeau, le large pantalon, la petite canne, la fuite, la chute, les glissades, l'équilibre instable, la soif, la faim, l'amour, chacune des particularités de Charlot avait sa voix propre, son intonation spéciale, son accent et son timbre. [...] Sa voix arrivait même à rendre la démarche les pieds en dehors de Charlot. »<sup>163</sup>

Monsieur Benjamin Nouvel-An, directeur d'un « cinéma-bistrot » dans lequel il officie en tant que bonimenteur, prête sa voix non seulement à Charlot – dont il prétend être, dans le programme reproduit dans l'ouvrage, le « sosie parlant »<sup>164</sup> – mais aussi à différents éléments de l'image. Sa voix colle à l'action ou à « l'atmosphère » (c'est-à-dire à la « psychologie » rudimentaire du burlesque) et contribue simultanément à individualiser les gestes et les attributs du personnage. Cendrars rend de façon poétique la représentation disloquée de l'individu induite par la segmentation résultant de la synchronisation vocogestuelle : chapeau, pantalon et canne semblent exister indépendamment de leur propriétaire, « animés » par la parole dont le rythme se calque sur celui de leurs déplacements. Ce n'est probablement pas un hasard s'il est justement question des films de Charlie Chaplin : la figure de Charlot représente, grâce à la « mécanisation » du jeu qu'implique sa précision pantomimique, l'exemple type de la décomposition des mouvements, voire de la fragmentation du corps – comme l'illustre le démembrement de la marionnette « cubiste » de Fernand Léger sur lequel s'ouvre le *Ballet mécanique*

163. Blaise Cendrars, *Dan Yack*, Paris, Denoël, 2002 [1<sup>re</sup> édition 1929], pp. 128-129.

164. *Ibid.*, p. 127.

(1924). L'adéquation entre la segmentation de la parole et l'image ne s'opère dès lors pas au niveau de la mise en chaîne, mais dans les rapports qui s'instaurent entre les paroles et les gestes visualisés à l'écran.

Ce texte dans lequel Cendrars dépeint de façon détaillée la pratique d'un bonimenteur prend une importance particulière en regard du contexte historique dans lequel il est produit: pour Cendrars, un personnage comme «Monsieur Nouvel-An» (qu'il ait vraiment existé ou non, puisque le réel se mélange de toute manière à l'imaginaire dans ces confessions fictionnelles qui s'apparentent à de l'autofiction)<sup>165</sup>, incarne «le véritable inventeur du cinéma parlant et du film sonore»<sup>166</sup>. En tant que grand manieur de mots, Blaise Cendrars fantasme dans sa description pétulante de la verve du bonimenteur une voix qui se fait geste, et la possibilité d'un rapport immédiat du public au langage qui permet de restituer tous les soubresauts de la vie. En fait, il inscrit cette apologie du boniment dans l'expression d'un idéal postulant que les «machines parlantes» sont capables de simuler l'«âme humaine». Il ne me semble pas fortuit que l'amie de Dan Yack, lorsqu'elle évoque ses expériences d'actrice de cinéma à travers le double intermédiaire de la voix de Dan Yack et de l'enregistrement phonographique, retrace avant tout le tournage d'un film qui n'est autre que l'adaptation du roman *L'Ève future* de Villiers de L'Isle-Adam – l'une des fictions les plus symptomatiques du paradigme culturel qui s'est constitué à partir de 1880 autour de l'imaginaire phonographique – qu'un cinéaste ironiquement nommé Lefauché (caricature d'Abel Gance) aurait réalisé au début des années 20. À l'instar de son héros, qui répète inlassablement ses enregistrements au dictaphone, Cendrars écrit en exergue de l'édition originale:

«Ce livre n'a pas été écrit. Il a été entièrement dicté au Dictaphone. Quel dommage que l'imprimerie ne puisse pas également enregistrer la voix de Dan Yack et quel dommage que les pages d'un livre ne soient pas encore sonores. Mais cela viendra. Travaillons.»<sup>167</sup>

165. Claude Leroy, en charge de la récente édition critique de *Dan Yack* parue chez Denoël, précise que Cendrars reprendra par la suite à plusieurs reprises le souvenir de «Monsieur Nouvel-An», notamment dans la revue *Montparnasse* et dans *Trop c'est trop*. Cette récurrence contribue à attester l'existence d'un substrat réel, ou du moins l'importance que cette pratique prit aux yeux de l'écrivain (B. Cendrars, *Dan Yack, op. cit.*, pp. 315-316, note 108).

166. *Ibid.*, p. 130.

167. *Ibid.*, p. 311.

En revenant au dictaphone qui avait été l'utilisation première à laquelle Edison destinait son phonographe, l'auteur révèle combien, en dépit de la mécanicité de l'enregistrement – qu'il fait ressentir à son lecteur en rendant par écrit les transformations phonétiques dues aux rayures des rouleaux – l'ère de la reproductibilité sonore correspond à un renouveau de l'oralité et de la vocalité. Or, comme en témoigne la digression sur « Monsieur Nouvel-An », le boniment synchronisé fait partie de ces nouveaux paradigmes représentationnels.

La synchronisation du boniment avec l'image que Blaise Cendrars attribue à l'exploitant d'une salle parisienne de la fin des années 10 (c'est-à-dire à une époque où le boniment a probablement disparu en tant que pratique courante, et fait donc office d'événement singulier) pouvait être à l'œuvre bien avant cette période. Par exemple, dans un article de 1901 relatant une prestation du bonimenteur québécois Henry Grandsaignes d'Haute-rièves, le chroniqueur précise que « chaque syllabe souligne un jeu de physionomie ou de scène »<sup>168</sup>.

Toute synchronisation n'était d'ailleurs pas non plus exclue dans les projections d'images fixes effectuées avec la lanterne magique. En effet, dans ce type de spectacle, le rythme de la parole interagit nécessairement avec la temporalisation induite par la succession des images (synchronisation de mise en chaîne), même si celle-ci n'est ni continue, ni gérée par des moyens mécaniques. Même atténuée par des procédés tels que le *dissolving views*, la scansion produite par chaque changement de plaque est incomparablement plus marquée que la transition d'un photogramme ou d'un plan à l'autre. Le principe de la mise en série d'images provoque l'introduction de seuils, c'est-à-dire de points potentiels de synchronisation avec le texte du commentaire. Cette ponctuation visuelle est accentuée en cas d'utilisation d'appareils à lanternes multiples (« polyoramas ») qui, selon Coissac, furent notamment « employés utilement pour simuler les apparitions, les orages et les éclairs »<sup>169</sup>. Les conférenciers tenaient compte de cette gestion temporelle, puisque leur préparation consistait notamment à prévoir les moments où le changement d'images devait avoir lieu, et à déterminer une longueur de texte appropriée à l'intérêt de l'image.

168. Cité par A. Gaudreault et G. Lacasse, « L'écran ventriloque », *art. cit.*, p. 139.

169. G.-M. Coissac, *op. cit.*, p. 14.

Dans les spectacles de lanterne où le nombre d'images est élevé, par exemple chez Alexander Black, les écarts entre les points de synchronisation sont réduits, offrant une synchronisation presque aussi continue qu'au cinéma, du moins du point de vue de la synchronisation de mise en chaîne. D'ailleurs, dans le «cinéma des premiers temps» où le «découpage» est fort sommaire (voire absent dans les films uniponctuels), les seuils correspondant aux changements de plans ne sont pas plus fréquents que lors d'une projection de lanterne magique. On rétorquera certes que l'activité de synchronisation du boniment de cinéma se réduit le plus souvent à joindre la parole et le geste, démarche impossible dans le cas de projections fixes. Il faut toutefois rappeler qu'il existait des lanternes – peu conservées jusqu'à aujourd'hui en raison de leur fragilité – qui permettaient, grâce à d'ingénieux mécanismes, certaines formes d'animation d'éléments représentés: les fantasmagores utilisaient la rétroprojection pour faire varier la grandeur des objets à l'écran en avançant ou reculant la source lumineuse dissimulée, ainsi que différents systèmes basés sur le principe de la superposition de verres<sup>170</sup>. Le caractère partiel, parfois discontinu et toujours ponctuel de tels mouvements les rendait plus aisément repérables que ceux de l'image cinématographique où l'on assiste à une infinité de mouvements visuels simultanés (sauf dans des films abstraits qui, justement, isolent certains mouvements). La discontinuité générale de la représentation (en dépit de tels mouvements continus) et l'absence de fixation temporelle du signifiant visuel ne fait donc pas obstacle à des effets de synchronisme vococinétique, et même vocogestuel lorsque l'image engage la représentation d'une figure humaine. C'est une synchronisation de ce second type que suggère Marcel Proust lorsqu'il évoque les «séances» où sa grand-tante manipulait pour lui une lanterne en animant le personnage de Golo tiré de la légende de Geneviève de Brabant:

«Golo s'arrêtait un instant pour écouter avec tristesse le boniment lu à haute voix par ma grand-tante et qu'il avait l'air de comprendre parfaitement, conformant son attitude avec une docilité qui n'excluait pas une certaine majesté aux indications du texte.»<sup>171</sup>

170. Voir L. Mannoni, *Le Grand Art... op. cit.*, p. 142.

171. Marcel Proust, *Du côté de chez Swann*, Paris: Le Livre de Poche, 1992 [1<sup>re</sup> édition 1913], p. 48.



Les traits humains attribués dans cette description au personnage représenté sur les plaques permettent à Proust de dépeindre le lampascope tel qu'il apparaissait aux yeux de l'enfant tout en révélant l'importance du rôle joué par le synchronisme dans cette perception anthropocentrique. On constate à la fois une coïncidence entre le verbal et l'image qui porte sur le référent (la conformité des attitudes de Golo) et une dimension temporelle (il s'arrête pour écouter) qui touche plus spécifiquement à la synchronisation. L'animation des personnages sur un fond ainsi que la succession des images permettent d'obtenir une synchronisation de mise en chaîne ou vococinétique. Le verbal «segmente» l'action, qu'il s'agisse des mouvements dans l'image ou de la dynamique créée par la mise en séquence. En retour, les ponctuations qu'introduit chaque plaque influent sur la perception du commentaire. Jérôme Prieur note à propos de ce passage de Proust l'influence du «rythme» de la succession des plaques sur l'écriture du romancier: «La langue mime le souvenir. Elle siffle, se hérissé, saute et sur-saute, progresse entre la nappe de lumière peinte et les secousses de la perception, plans exorbités qu'interrompent de brèves fermetures au noir.»<sup>172</sup> La réminiscence figurée par Proust, qui convertit une perception audiovisuelle en langage verbal, vise en effet à «mimer» le boniment, comme le texte de Prieur mime à son tour le «phrasé» proustien. Les expressions utilisées par Prieur relèvent de l'isotopie de la «saccade» (sautes, secousses et brèves interruptions) que l'on trouve effectivement chez Proust, notamment dans les expressions «au pas saccadé de son cheval» et «s'avancait en tressautant»<sup>173</sup>. L'exemple de l'évocation proustienne de séances de projections domestiques avec lanterne montre qu'une perception basée en partie sur le synchronisme pouvait être convoquée dans des spectacles pré-cinématographiques où l'image était «fixe». La voix occupait donc une place déterminante dans ce genre de représentation. Même lorsqu'elle n'intervenait qu'à titre d'intermède, elle permettait de définir le rythme de la succession des vues et celui du déroulement général de la «séance», créant à ces deux niveaux une synchronisation de mise en chaîne.

172. J. Prieur, *Séance...*, *op. cit.*, p. 133.

173. M. Proust, *Du côté de chez Swann*, *op. cit.*, pp. 47-48.



## II. LA MEDIATION BONIMENTORIELLE

### II.1 LE BONIMENT COMME PERFORMANCE

À l'instar d'autres émissions *live* (musique d'accompagnement, effets sonores), le bonimenteur instaure une *relation de proximité* particulière avec l'audiospectateur. Proximité de *nature* dans le sens où il s'agit d'une instance humaine qui s'exprime par le biais d'un canal spécifiquement humain; proximité *physique* puisqu'il est situé dans le même espace-temps que son auditoire, plus précisément «à portée de voix». Dans sa description d'une séance de lanterne magique, G.-M. Coissac souligne l'importance de la relation privilégiée que l'orateur entretient avec son public:

«[...] le conférencier se rend compte, par l'attitude de son auditoire, par son attention, ses émotions, s'il a compris et s'il ne parle pas en vain. Ce n'est plus une voix qui se fait entendre, c'est une âme vivante, qui vibre et fait vibrer, une âme qui parle à des âmes.»<sup>174</sup>

Le lyrisme qui point soudain dans ce manuel technique révèle combien, pour cet auteur rompu à l'exercice des projections-conférences, l'intervention orale véhicule une profonde humanisation. En dépit de l'origine mécanique de l'image, le spectacle est empreint, grâce au boniment, d'une véritable sensibilité; la voix *anime* la représentation. Comme on l'a vu à propos du spectacle de lanterne magique, le locuteur capte l'attention des spectateurs et l'oriente plus ou moins explicitement vers l'un des deux autres pôles du dispositif. Il existe donc deux types de médiation bonimentorielle: celle qui définit le rapport du spectateur à la machinerie,

174. G.-M. Coissac, *op. cit.*, p. 179.

et celle qui régit le rapport de celui-ci à la représentation. À l'intérieur de cette configuration, on peut envisager une double orientation de la médiation bonimentorielle (spectateur ↔ représentation ; spectateur ↔ machinerie) à partir du noyau que constitue la relation au spectateur.

Le fort lien de proximité établi avec l'(audio)spectateur résulte de la fonction même de la « médiation », manifeste dans le sentiment de l'*unicité* du spectacle qui conduit à la *communion* de ses participants. Si l'*unique* contribue à unir, c'est parce qu'il manifeste le partage, d'autant plus important qu'il est éphémère, de quelque chose de *commun*. Le spectateur a véritablement l'impression d'« appartenir » au spectacle du fait qu'il sait qu'en dehors de sa propre expérience présente, la performance ne peut se répéter à l'identique. Serge Cardinal, reformulant une remarque de Miriam Hansen sur la présence d'accompagnements vivants durant les projections des premiers temps, résume ce qui fait l'importance de la dimension sonore dans la constitution de l'expérience de l'audiospectateur du cinéma « muet » :

« Autrement dit, la projection n'a ce caractère toujours singulier de la performance qu'à travers l'activité, à chaque fois reprise, des productions sonores ; le public y trouve une présence à soi toujours renouvelée et spécifique et tient son expérience moins du film que de l'interprétation sonore qui l'a enveloppé. »<sup>175</sup>

L'auteur utilise l'expression de « performance » pour rendre compte de la particularité des productions sonores. Ce terme évoque les travaux de Paul Zumthor, qui axe son étude de la poésie orale sur cette notion empruntée aux folkloristes et ethnologues. Il définit la « performance » comme « l'action complexe par laquelle un message poétique est simultanément transmis et perçu, ici et maintenant »<sup>176</sup>. Cette approche dynamique de la communication orale s'applique également au boniment des vues animées, indépendamment de la « poéticité » du texte. C'est pourquoi j'aborderai cette pratique à travers certains aspects envisagés par Zumthor dans *Introduction à la poésie orale*.

175. Serge Cardinal, « Médiation ou modulation sonore? », *Cinémas*, Vol. 9, N° 1, 1999, pp. 97-98.

176. P. Zumthor, *Introduction...*, *op. cit.*, p. 32.

L'audiospectateur du « muet » est guidé par le bonimenteur qu'il appréhende sur le mode de la *tactilité* que Zumthor attribue à tout énonciateur d'une parole vive: «L'oralité implique tout ce qui, en nous, s'adresse à l'autre: fût-ce un geste muet, un regard.»<sup>177</sup> Le bonimenteur interpelle en effet le spectateur dans une démarche active de canalisation du regard et de l'écoute, parfois si concrète qu'elle s'apparente à une préhension. Ainsi pourrait-on dire, pour utiliser trois termes possédant la même racine, que le bonimenteur *capte* (l'ouïe, le regard, l'attention), *captive* (en soulignant le caractère vivant du spectacle, notamment par l'exploitation du potentiel expressif de la voix), voire *capture* le spectateur (dans les rets de ses explications, de son récit, etc.). Cette triple fonction pourrait nous conduire à conférer au bonimenteur une dimension « disciplinaire » telle que la conçoit Noël Burch qui, dans un article ultérieur à ses premiers commentaires sur le *benshi* et le conférencier, considère le boniment comme une activité « policière et éducative »<sup>178</sup>. En tant qu'adjuvant de la narration et instance interprétative, le locuteur a tendance, il est vrai, à agir de façon coercitive sur la lecture personnelle du spectateur. En 1906, un rédacteur de la *Phono-Ciné-Gazette* s'offusquait en ces termes de l'intervention envahissante du bonimenteur:

«[...] la scène, si elle est bien mimée, ne doit pas avoir besoin d'explications verbales ou écrites. Je hais le «bonisseur» qui me prend, moi, public, pour un imbécile incapable de comprendre ce qu'on lui fait voir et ne me laisse pas le plaisir de deviner.»<sup>179</sup>

Au vu du caractère rudimentaire du travail filmographique effectué sur les films de l'époque, l'expression pantomimique du profilmique devait convoquer la plus grande partie de la signification. Si l'auteur de l'article joue l'image contre le verbe, c'est avant tout pour affirmer son droit à interpréter lui-même les actions montrées, à ne pas être dépossédé de cette activité cognitive qui est, selon lui, source de « plaisir ». Cette opinion s'inscrit tout à fait dans le paradigme oppositionnel discipline *vs.* jouissance dont Noël Burch fait l'hypothèse. Toutefois, les conditions de l'oralité posent des limites au pouvoir coercitif du verbal, qui par ailleurs

177. *Ibid.*, p. 193.

178. N. Burch, «Discipline et jouissance», *art. cit.*, p. 91.

179. *Phono-Ciné-Gazette*, N° 35, 1<sup>er</sup> septembre 1906, cité in J. Meusy, *op. cit.*, p. 140.

varie selon les usages (spectacle forain, projection à visée pédagogique, etc.). Inscrit dans une performance, le boniment est indissociable des réactions de son auditoire qui, gagné par le caractère « vivant » de la présentation, ne peut être confiné à une totale passivité, puisqu'il est convié à une dynamique collective de participation.

On trouve dans un témoignage du journaliste et critique Isroel Rabon (*La Rue*, 1928), repris dans *Le Spectateur nocturne* de Jérôme Prieur, une illustration de l'entrain contagieux que la voix était capable d'insuffler aux spectateurs à l'époque du « cinéma » parlé. La première voix dont le texte fait mention, pure « voix-attraction », est celle du patron de l'établissement qui, « emballé par sa propre éloquence », fait les louanges de la performance à venir de l'orateur<sup>180</sup>. Il crée ainsi, par un discours antérieur à la projection, un horizon d'attente lié à « l'attraction » que représente le boniment proprement dit. Après que Rabon a pris la parole pour commenter un film historique (non identifié, mais dont on comprend qu'il porte sur la Révolution française), la vivacité du bonimenteur gagne le public prolétaire de la salle. Rabon déclare : « Mes paroles avaient enthousiasmé les ouvriers chauffés à blanc par la perspective de la grève générale. Un brouhaha indescriptible régnait dans la salle. » Indépendamment du contexte historique posé par la diégèse du film, le public a donc fait sienne la situation montrée en l'interprétant par rapport à ses préoccupations du moment. Il se peut bien, comme l'indique l'auteur, que cette exaltation collective ait été favorisée par le contexte de la performance orale, qui démontre dans ce cas certaines vertus soulignées par Zumthor. À la fois ancré dans un présent qui est aussi celui de ses destinataires et proféré dans une situation de spectacle favorable à la dissolution de l'individualité, le boniment, parole vive, déclenche des réactions houleuses. Les spectateurs *se manifestent*, la séance vire à la *manifestation*. C'est alors que le rapport de « coercition » entre espace de production et spectateur s'inverse : ce sont les spectateurs qui exigent de l'orchestre qu'il joue *La Marseillaise* et *Drapeau rouge*, puis contraignent le projectionniste à recommencer la projection. Le passage à la chanson représente, comme l'a souvent noté Zumthor, le point culminant de l'expression de l'oralité ;

180. Cité par Jérôme Prieur, *Le Spectateur nocturne*, Paris : Éditions de l'Étoile, Cahiers du cinéma, 1993, p. 89.

ici, les airs entonnés collectivement contribuent à la consolidation du groupe dans une effusion spontanée qui est véritablement de l'ordre de la *communio*n. Il n'est en ce sens pas étonnant que l'enseignement religieux ait exploité le pré-cinéma et le « cinéma » parlés pour rassembler les fidèles. Ces spectacles rejouent en effet, à travers le maintien d'une relation de proximité, les rituels chrétiens de profession d'une foi commune.

Même si ce fait relaté par Rabon est assez exceptionnel, il illustre le potentiel d'aléatoire dont la parole peut disposer dans un spectacle passablement réglé par des contingences mécaniques. Pour Zumthor, les auditeurs sont partie intégrante de la performance, et cela même au cinéma où, en dépit des conditions d'isolement du spectateur, « peut se former (dans une salle assez bien remplie) une sensation confuse d'unanimité, apte à provoquer parfois une réaction commune »<sup>181</sup>. Le récit de Rabon montre que *tout* boniment ne conduit pas, ainsi que semble le signifier Noël Burch, à régenter le rituel de la séance en favorisant la passivité.

On peut cerner un peu plus précisément la nature de la médiation bonimentorielle en se référant à l'étude de la performance orale proposée par Zumthor. Il discute en effet certaines composantes qui me paraissent similaires aux caractéristiques de la « situation d'énonciation » telles que l'envisagent communément les linguistes depuis les travaux d'Émile Benveniste. Ainsi toute performance implique-t-elle des acteurs (émetteurs et récepteurs), des moyens (la voix, le geste et des « médiats ») ainsi que des circonstances temporelles et spatiales<sup>182</sup>. Ces dernières ont trait à la présence d'un décor auquel est associée non seulement la vue, mais un ensemble sensoriel. Pour le boniment, il s'agit de la salle de cinéma : il est évident que l'atmosphère qui y règne pèse sur la réception de la parole. Zumthor remarque qu'il est inévitable que la performance soit « parasitée de « bruits »<sup>183</sup>. Or, comme l'affirment certains textes d'époque (voir II.4), le bonimenteur était nécessairement confronté au bruit du projecteur et au brouhaha du public.

En utilisant le terme d'« acteur » et en évoquant fréquemment la question de la théâtralité, le médiéviste met également l'accent sur la dimension visuelle de la poésie orale. Cet aspect ne doit pas être

181. P. Zumthor, « Le geste et la voix », *art. cit.*, p. 74.

182. P. Zumthor, *Introduction...*, *op. cit.*, p. 149.

183. *Ibid.*, p. 156.

négligé dans une prise en compte théorique du bonimenteur: son habillement, sa mimique, sa gestuelle pouvaient contribuer grandement à l'orientation du sens donné aux images. Certes, le locuteur n'était pas toujours visible durant la projection, mais peut-être l'est-il avant ou après. Entre le pôle de sa mise en spectacle quasi scénique et celui de son effacement dans l'obscurité qui baigne l'espace spectatoriel, quantité de positions médianes pouvaient se présenter qui contribuaient à créer une représentation visuelle *live* (par opposition à celle, *écranique*, que livre le film). Ainsi la bipartition image/parole, commode lorsque l'on considère rétrospectivement le «cinéma» parlé à partir d'un support qui fait se côtoyer piste-son et bande-image, ne tient plus: la représentation visuelle elle-même se voit dédoublée, ses deux facettes interagissant selon diverses modalités (complémentarité, redondance, opposition). L'espace de la salle est occupé par trois vecteurs de représentation: la voix, le corps et l'image projetée. Alors que le dernier d'entre eux nécessite un processus d'intégration à la performance, les deux premiers sont intimement liés, presque jusqu'à se confondre. C'est pourquoi, même soustrait aux regards, le bonimenteur propage une voix qui est perçue comme la trace de sa corporalité. La force de sa présence, actualisée à travers une voix qui, comparativement au texte écrit, se manifeste comme «toute-présente»<sup>184</sup>, ancre l'image dans l'espace-temps performantiel sur le mode de l'immédiateté. Véritable participant de l'événement, l'audiospectateur est happé dans le dispositif.

L'effet de présence induit par les pouvoirs de la vocalité interdit selon moi de considérer le boniment comme opérant *fondamentalement* une distance avec la représentation, ainsi que le suggère André Gaudreault en disant que «la distance créée par le commentaire du bonimenteur entre le monde diégétique mis sur le film et le spectateur reste vraisemblablement irréductible»<sup>185</sup>. Parmi les différentes postures pour lesquelles le bonimenteur était en mesure d'opter, la mise à distance de l'univers diégétique constitue selon moi une composante facultative, et non une donnée consubstantielle à ce type de spectacle. En outre, si nous avons une connaissance historique plus précise de ces phénomènes, nous découvririons probablement que des changements de posture survinrent

184. *Ibid.*, p. 234.

185. A. Gaudreault, «Fonctions et origines...», *art. cit.*, p. 143.



avec l'autonomisation de la diégèse et de la narration qui s'opère autour de 1910. Cependant, l'absence d'intégration du discours bonimentoriel au «texte» filmique ne constitue pas nécessairement un facteur de mise à distance, dans la mesure où cette extériorité accroît la dimension participative du spectacle. L'engagement corporel et vocal du bonimenteur reconduisant un type de transmission qui a fait ses preuves, sous diverses formes, depuis des siècles, le boniment a également pu favoriser l'immersion du spectateur dans la représentation. La fonction de médiation entraîne un renforcement du rapport d'intimité entre l'espace de la production et le spectateur qui, plus généralement, caractérise toute performance. D'ailleurs, une projection muette d'un film des premiers temps n'était pas plus propice à l'immersion du spectateur que son équivalent bonimenté. Il faut veiller à ne pas attribuer à la seule présence du bonimenteur des effets qui appartiennent en fait aux conditions générales de la réception du cinéma des premiers temps.

En tant qu'instance narrative, le bonimenteur se met au diapason des spectateurs, ainsi que le note R.-M. Arlaud à propos de Louis Régnault: «Les histoires sont simples. Lorsqu'elles se compliquent, il les refabrique à son idée ou selon son public, car lui est psychologue et seul eut le droit de dire: *Mon public.*»<sup>186</sup> L'ingérence bonimentorielles dans l'organisation narrative a donc plutôt pour effet de créer une proximité entre le bonimenteur et son public. Le terme «distance» ne me semble pas se départir de sa signification littérale liée à l'espace, soit ici à la disposition matérielle du dispositif, et cela même dans son sens figuré. Or le bonimenteur fait le plus souvent partie de l'espace physique occupé par le public, partageant avec ce dernier une même corporalité dont la voix, englobante et unificatrice, est l'émanation privilégiée, alors que les images bidimensionnelles en sont fondamentalement dépourvues. Je ne prétends pas que l'insertion d'un locuteur dans le dispositif cinématographique n'affecte en rien le statut de la voix, ni les processus de réception induits par ledit dispositif. Il me semble cependant que c'est faire trop grand cas de la seule représentation visuelle écranique que de concevoir le dispositif du «cinéma» parlé comme «ontologiquement» défini par une distance. On ne peut nier à la parole une capacité d'introduire le

186. R.-M. Arlaud, *op. cit.*, p. 76.

spectateur dans un univers fictionnel, comme le firent, entre autres conteurs oraux, les troubadours médiévaux ou les griots africains. Pourquoi la coprésence du conteur et d'une image projetée enrayerait-elle cette faculté ancestrale de la communication orale?

Il ne faut pas sous-estimer le pouvoir de présentification de la voix. On peut à mon sens postuler que les images, actualisées à travers la voix vive, héritent également de son « actualité ». Certes, la nature photographique de l'image cinématographique implique, ainsi que le souligne Gaudreault lorsqu'il décrit l'instance bonimentorielle comme « posée en retrait d'un récit qui s'est produit « ailleurs » et « avant », un autre temps et un autre espace que ceux qui caractérisent la performance vivante<sup>187</sup>. Toutefois, le « récit » s'actualise majoritairement dans le présent de l'énonciation verbale, et l'espace-temps originaire se voit réintégré en tant que projection dans le *hic et nunc* du spectacle. L'univers diégétique est issu de la *coréférence verbo-iconique* – le verbal du boniment se joint à l'iconique de l'image – et n'est pas assimilable au moment du tournage des bandes. Ce processus de présentification par la parole correspond également au devenir du texte écrit dans la performance orale, dont Paul Zumthor rend compte en évoquant une « suspension des effets de distance historique entre la genèse de l'œuvre poétique et chacune de ses réalisations »<sup>188</sup>. Cette abolition de la dimension temporelle concerne aussi, dans une certaine mesure, l'écart qui sépare la genèse du film de sa « réalisation » (c'est-à-dire du moment où il devient « réel » pour un spectateur) au sein du spectacle bonimenté. Certes, la pratique de la poésie orale se substitue à l'écrit, alors que, bien souvent, le boniment est proféré simultanément à la projection d'images. Cependant, placée sous l'autorité du verbal, la projection apparaît dans le présent de son déroulement.

La question de la « distance » constitue un aspect central de la théorisation des liens entre image et voix, et cela également dans le cas du cinéma parlant. Même lorsqu'il privilégie l'adresse au spectateur, le boniment ne peut être mis sur le même plan que la « confrontation exhibitionniste » de l'image. Le lien qu'il établit avec le public ne résulte pas de la construction d'une « diégèse ». On retrouve là l'idée que Noël Burch se faisait de la fonction du

187. A. Gaudreault, « Fonctions et origines... », *art. cit.*, p. 143.

188. P. Zumthor, *La Lettre...*, *op. cit.*, p. 115.

*benshi*, elle-même influencée, me semble-t-il, par la conception brechtienne de la «distanciation». Le point de vue de Burch sur la question, largement informé par des postulats esthétiques formulés à l'encontre du cinéma contemporain (dans la mouvance des années 70 d'une critique idéologique du mode dominant de représentation), consiste à considérer le *benshi* comme un facteur d'enrayement de la fonction illusionniste de la représentation. Par la suite, Burch opposera à cette mise à distance l'effet inverse du boniment occidental, qui aurait prioritairement contribué à la linéarisation du récit. Or, ainsi que l'a proposé Tom Gunning<sup>189</sup>, il est possible d'envisager ces deux fonctions comme deux pôles extrêmes de l'activité du bonimenteur. Il s'agit donc, répétons-le, du choix d'une posture énonciative plutôt que d'une donnée intrinsèque à la projection bonimentée. En faisant de l'extériorité «diégétique» fondamentale du bonimenteur un critère décisif (avec cette idée qu'il «injecte» quelque chose dans un autre moyen d'expression), on court le risque de réintroduire une perspective textuelle dans la prise en compte d'une pratique spectaculaire au lieu d'élargir le «film» à son contexte institutionnel de réception. Le critère de l'extériorité est par contre valable si l'on compare cinéma parlé et cinéma parlant, puisqu'il est nécessaire de distinguer dans ce cas la médiation concrète du bonimenteur et la production d'une voix-over par la machinerie. L'intervention bonimentorielle apparaît comme singulière au sein de l'environnement qui compose l'instance de production de la représentation. Si, comme l'a constaté Rick Altman, un conférencier tel que Lyman Howe abandonne après quelque temps les effets enregistrés qui exigent de briser l'illusion en manipulant un phonographe<sup>190</sup>, ce n'est pas tant pour accroître le pouvoir illusionniste de la représentation que pour assurer la domination d'une gestion *humaine* de l'espace de production. L'immersion du spectateur s'appuie moins sur la «suspension volontaire de l'incrédulité» que sur l'attitude participative à laquelle engage toute pratique vivante associée à la «performance» telle que la conçoit Paul Zumthor.

189. T. Gunning, «The Scene of Speaking...», *art. cit.*, p. 71.

190. R. Altman, *Silent Film Sound*, *op. cit.*, p. 147.

## II.2 BONIMENT ET « DISCOURS INTERIEUR »

À une époque antérieure à la généralisation du « polissage » de la consommation des films, on peut imaginer que les baraques foraines étaient fréquentées par une assemblée peu discrète. Ce climat bruyant a pu se perpétuer lors de certaines projections plus tardives. Le cinéaste Roger Lion, soucieux de souligner combien l'exploitation cinématographique française souffre en régions rurales d'importantes carences (notamment en l'absence de toute musique d'accompagnement), décrit ainsi une séance où son film *Les Fiançailles rouges* (1926) fut projeté par un exploitant ambulant :

« Les pellicules, aux perforations arrachées, aux scènes mille fois recollées et à demi dévorées, n'offraient plus qu'une histoire invraisemblable, dont la salle entière épelait à haute voix les intertitres brefs, dominant pour un instant le bruit infernal de la boîte à musique du carrousel voisin. »<sup>191</sup>

Dans ce contexte assourdissant qui rappelle l'exploitation cinématographique des premières années du siècle, les spectateurs sont incités à se manifester verbalement. Lorsque la dimension narrative de l'organisation filmique devient secondaire – situation comparable à l'absence de narration intrinsèque dans le cinéma des premiers temps où dominaient les attractions – l'emprise du film sur le public se relâche. La diminution du caractère coercitif du spectacle est favorable à l'expression du public par la voix vive, ne serait-ce que pour ânonner les quelques bribes de récit perceptibles dans une tentative collective d'attribution du sens. La voix s'oppose par conséquent aux conditions d'isolement du régime de perception dominant au cinéma que Boris Eikhenbaum décrivait ainsi :

« L'état d'âme du spectateur est proche de la contemplation solitaire, intime; il regarde en quelque sorte le rêve d'un autre. Le moindre bruit étranger au film l'irrite beaucoup plus qu'au théâtre. Le bavardage des voisins (la lecture à haute voix des intertitres, par exemple) l'empêche de se concentrer sur le mouvement

191. Roger Lion, « En avant la musique! », *Cinémagazine*, N° 47, 19 novembre 1926, p. 398 (je souligne).

du film; l'idéal serait de ne pas sentir la présence des autres spectateurs, d'être en tête-à-tête avec le film, de se sentir sourd-muet.»<sup>192</sup>

Le théoricien soviétique envisage un état du cinéma qui s'est imposé depuis la seconde moitié des années 10 et qui se caractérise par un «discours» visuel autosuffisant. Les images s'offrent au spectateur dans un rapport d'immédiateté qui écarte toute dimension communautaire. Cependant, des manifestations sonores comme celle qu'évoque Roger Lion viennent rompre cette intimité de «sourd-muet» pour imposer la présence collective de la masse. Cette situation singulière révèle par contraste combien les sons issus de «l'espace de production», de quelque nature qu'ils soient, sont susceptibles de contraindre le public au silence, et donc à une intériorisation de la parole: les intertitres, lorsqu'ils sont lus «mentalement», perdent l'effet de présentification de l'oralité pour s'accorder à la présence-absence des êtres de l'écran. On comprend que certains promoteurs du début du siècle les supprimaient afin qu'ils n'interfèrent pas avec le boniment, ainsi que l'indique Arlaud: «Lorsque Pathé sous-titrera ses films, les ambulants se vexeront d'une dignité qui leur échappe, d'ailleurs ils couperont les morceaux de texte afin de continuer à <faire la conférence.>»<sup>193</sup> Dans le cas mentionné par Lion, l'oralité s'exerce toutefois dans la salle, de sorte qu'elle échappe à la gestion du spectacle prévue par l'exploitant. Le réalisateur fait d'ailleurs mention d'une autre projection durant laquelle le directeur eut recours à un phonographe qu'il fit jouer en même temps que le film pour «faire taire le monde»<sup>194</sup>. La musique, le bonimenteur ou les machines parlantes constituent un facteur de réglementation visant à transformer fictivement les «sons de la salle» en «sons du film». Ce rôle imparti aux sons était encore plus important à l'époque des premiers exploitants sédentaires, lorsque le public était plus indiscipliné. Dans un article de 1909 significativement intitulé «La parole humaine comme facteur dans le spectacle de vues animées», W. Stephen Bush, collaborateur du journal spécialisé *Motion Picture World*, qui exerçait parallèlement la profession de bonimenteur, soulevait le problème suivant:

192. Boris Eikhenbaum, «Problèmes de ciné-stylistique» [1927], in F. Albera (dir.), *Les Formalistes russes et le cinéma. Poétique du film*, Paris: Nathan, 1996, p. 43.

193. R.-M. Arlaud, *op. cit.*, p. 76.

194. R. Lion, «En avant la musique!», *art. cit.*, p. 398.

«Prenez place parmi le public et qu’observez-vous? Lorsque l’histoire avance, et même à ses tout débuts, ceux qui sont doués d’un peu d’imagination et d’éloquence commenceront à commenter, à parler avec plus ou moins d’agitation et à essayer d’expliquer certaines choses à leurs amis et voisins. Ce courant d’électricité mentale grandira et deviendra sauvage, irrégulier, incontrôlable.»<sup>195</sup>

Pour remédier à cette absence d’emprise de l’espace de la production sur les spectateurs, et supprimer l’angoisse suscitée par une situation impossible à maîtriser (ainsi que le suggèrent les trois derniers épithètes de la citation), Bush propose d’utiliser un bonimenteur qui saura canaliser ce désir d’expression «en se faisant sans effort le *porte-parole* de la foule particulière d’êtres humains qui constituent le public»<sup>196</sup>. Cette formulation indique clairement que le bonimenteur se substituait, en tant qu’instance, au spectateur «parlant».

À l’époque où Roger Lion publie l’article discuté ci-dessus (au milieu des années 20), la parole a été totalement intégrée au film sous la forme des intertitres. Les voix se sont tues, et il incombe désormais à la musique d’imposer le silence dans la salle. Partisan, comme la plupart des cinéastes<sup>197</sup>, d’un contrôle de l’instance de production sur la réception du film, Roger Lion prône l’accompagnement musical. On constate que pour Boris Eikhenbaum – dont la théorie me paraît intéressante à convoquer dans ce contexte – la musique d’écran constitue justement une composante favorable à la formation de ce qu’il appelle le «discours intérieur»<sup>198</sup>. Cette notion permet à mon avis de définir *a contrario* la fonction de médiation du bonimenteur, qui actualise oralement un «langage» interne au film pour l’adresser à un destinataire collectif. Dans ce sens, on peut considérer que le boniment véhicule des contenus et fournit les codes de lecture du film. Se référant à la dichotomie proposée par Piaget et rediscutée par Lev Vygotski<sup>199</sup>,

195. Cité in R. Altman, *Silent Film Sound*, *op. cit.*, p. 279.

196. *Ibid.*, p. 280.

197. R. Altman, «Naissance de la réception classique...», *art. cit.*, p. 106.

198. B. Eikhenbaum, «Problèmes de ciné-stylistique», *art. cit.*, p. 47.

199. Alors que le langage égocentrique constitue pour Piaget une forme primaire de satisfaction des besoins sans prise en compte de la réalité, et donc un stade intermédiaire de l’évolution de l’individu vers un langage intelligent (qui se mettrait en place à l’âge de 8 ans environ), Vygotski démontre sur la base d’une analyse expérimentale que le langage égocentrique s’apparente également à la logique de l’action, et soutient la thèse selon laquelle il se perpétue chez l’adulte. Voir Lev Semenovic Vygotski, *Pensée et langage*, Paris: Éditions Sociales, 1985, pp. 95-103.

François Albera rend compte de la particularité de l'hypothèse théorique du « discours intérieur » envisagée par Eikhenbaum en affirmant qu'elle consiste à « offrir un intermédiaire entre le discours « égocentrique » et le discours socialisé »<sup>200</sup>. Or le bonimenteur, facteur d'extériorisation du discours du film (il transforme en quelque sorte l'« endophonie » en phonie), se situe indéniablement du côté de la socialisation sans toutefois l'effectuer sur le mode conventionnel de l'écrit (à l'instar des livrets qui pouvaient accompagner la projection d'un film à l'époque muette). Lorsque Vygotski illustre la parenté profonde du langage égocentrique de l'enfant – qu'il considère, contrairement à Piaget, comme relevant déjà du social, mais qu'il oppose au langage « communicatif » – et la réflexion silencieuse de l'adulte, il fait référence aux expériences qui ont consisté à faire résoudre à un adulte un problème mental à haute voix<sup>201</sup>. L'oralité joue donc un rôle décisif dans la prise de conscience de ces phénomènes de verbalisation non soumis aux impératifs de la transmission à autrui. En prêtant sa voix aux réflexions muettes suscitées par l'image animée tout en les modifiant pour que les significations élaborées au cours de ce processus soient transmissibles, le bonimenteur effectue un double mouvement de socialisation, dont la portée est accentuée par l'adresse à un auditoire collectif.

Chez Eikhenbaum, cette question du langage intérieur qui ressortit globalement à la psychologie trouve une application directe dans le cadre des conditions de réception d'un film. En examinant quel travail cognitif Eikhenbaum impute au spectateur de cinéma, on peut dès lors tenter de cerner plus précisément les fondements de l'activité bonimentorielle. Le « discours intérieur » nous conduit à penser que la verbalisation produite par le bonimenteur ne se limite pas à familiariser le public avec le monde du film, dont l'appréhension est balisée par la désignation verbale de certains éléments nodaux, mais aussi à lui faire comprendre l'organisation signifiante propre au cinéma. Le concept de « discours intérieur », pierre angulaire de la réflexion d'Eikhenbaum dans la mesure où il permet de légitimer le recours à des équivalences entre le cinéma et les langues naturelles<sup>202</sup>, est en effet lié à l'organisation syntagmatique du film. De ce fait, les implications sont d'ordre sémantique,

200. François Albera, « Introduction », in *Les Formalistes russes...*, *op. cit.*, p. 19.

201. L. Vygotski, *op. cit.*, pp. 100 et 105.

202. B. Eikhenbaum, « Problèmes de ciné-stylistique », *art. cit.*, pp. 51-52.

logique, temporel et plus largement narratif. Boris Eikhenbaum prend en considération le mouvement progressif de compréhension qu'opère mentalement le spectateur afin de construire une unité à partir de fragments (les «ciné-phrases»), puis d'atteindre ce niveau supérieur d'articulation que sont les «ciné-périodes». Il me semble possible de rapporter cette conception du discours intérieur, basée sur l'établissement d'une continuité temporelle et d'une cohérence, à la fonction du bonimenteur lorsque ce dernier visait à clarifier le contexte narratif d'un film :

«Chaque scène est présentée au spectateur par fragments, par à-coups. Beaucoup de choses restent hors de sa vision : les intervalles entre ces à-coups sont remplis par le discours intérieur.»<sup>203</sup>

On peut en effet supposer que le boniment a contribué de la même façon à combler les manques du récit, exhibant ainsi les conventions qui régissent les opérations de «montage». C'est pourquoi il a sans doute joué un rôle considérable dans le développement de certaines compétences de lecture chez le spectateur du cinéma des premiers temps. Lorsque Eikhenbaum explique que les réalisateurs doivent faire en sorte que le spectateur «devine le sens de l'épisode ou, en d'autres termes, qu'il le *traduise* dans le langage de son discours intérieur»<sup>204</sup>, on retrouve un principe, celui de la traduction, qui me paraît être, dans la multiplicité des niveaux qu'il engage, au cœur de l'activité bonimentorielle. Emilio Garroni, discutant l'hypothèse du Soviétique à un niveau plus général (celui du «langage interne», qui l'amène à s'interroger sur les correspondances existant entre pensée et langage), fait remarquer que «nous pouvons avoir de la difficulté à traduire un discours interne en discours externe, *comme si* la différence entre les codes respectifs faisait naître des problèmes et des difficultés de traduction analogues à ceux de la traduction proprement dite»<sup>205</sup>. Garroni s'oppose à cette conception car il associe la notion d'Eikhenbaum à celle de «langage intérieur». Il faut cependant noter que le «discours intérieur» s'en distingue dans la mesure où il est suscité par une série de paramètres extérieurs précis que sont les images du film.

203. *Ibid.*, p. 59.

204. *Ibid.*, p. 45.

205. Emilio Garroni, «Langage verbal et éléments non-verbaux dans le message filmico-télévisuel», in Dominique Noguez (dir.), *Cinéma : théorie, lectures*, Paris : Klincksieck, 1973, p. 116.



L'audiospectateur n'est pas totalement coupé du monde comme il l'est dans le cas de l'endophonie, mais il interagit avec les pôles du dispositif cinématographique. Ce distinguo transparait d'ailleurs dans l'utilisation française du terme « discours » : il s'agit d'une actualisation particulière d'un « langage » (en fait la rencontre de la pensée du spectateur et du « langage » filmique), conditionnée par l'organisation du film. On comprend donc que le modèle « sémio-cognitivist » du discours intérieur envisagé par Eikhenbaum n'exclut pas la prise en compte des phénomènes d'extériorisation. En ce sens, on peut admettre que le boniment tend à se substituer au discours intérieur, dont il court-circuite l'exercice en agissant comme une médiation concrètement intégrée au dispositif. C'est pourquoi Eikhenbaum, abordant la question qui lui était familière du tirage des films muets<sup>206</sup>, prône le recours aux intertitres de dialogue tout en s'opposant farouchement aux mentions verbales extradiégétiques. En effet, les intertitres explicatifs<sup>207</sup> ou narratifs<sup>208</sup> brisent selon lui le discours intérieur du spectateur en introduisant ce que l'on appellerait aujourd'hui une « instance énonciative », qu'il nomme « auteur » ou « auteur-narrateur »<sup>209</sup>. Or cette origine du verbal me paraît comparable au bonimenteur en ce qu'elle médiatise explicitement, par la parole, le rapport du spectateur au film.

En s'extériorisant à travers le boniment, le discours intérieur migre du spectateur vers une instance de médiation, et donne lieu à une verbalisation effective, alors que, selon la conception d'Eikhenbaum, ce discours relève d'une opération de la pensée qui est à la fois globalement conditionnée par les catégories du langage verbal et située à un niveau pré-linguistique. Le théoricien soviétique précise à cet égard que « le discours intérieur du spectateur est beaucoup plus fluide et vague que le discours proféré »<sup>210</sup>. La parole du bonimenteur introduit par contre une actualisation spécifique d'un certain nombre de sèmes présents dans l'image. En outre, lorsque le bonimenteur s'exprime en discours direct afin de prêter sa voix aux personnages de l'écran, il impose certaines caractéristiques vocales

206. Voir Alexander Schwarz, *Der Geschriebene Film. Drehbücher des deutschen und russischen Stummfilms*, Munich: Éditions Diskurs Film, 1994, p. 90.

207. B. Eikhenbaum, « Le mot et le cinéma », in F. Albera (dir.), *Les Formalistes russes...*, *op. cit.*, p. 210.

208. B. Eikhenbaum, « Problèmes de ciné-stylistique », *art. cit.*, p. 45.

209. *Idem.*

210. *Ibid.*, p. 46.

qui demeurent virtuelles lors d'une projection muette. L'une des craintes qui apparaît de façon récurrente dans les discours contemporains des premiers *talkies* est celle de voir disparaître le caractère universel du cinéma « muet » (non bonimenté, mais accompagné de musique). Jean Keim explique ainsi les changements auxquels fut confronté public de cinéma de la fin des années 20 :

« Dans le silence, ou plutôt dans l'atmosphère musicale appropriée, il avait l'habitude de faire prononcer aux acteurs les mots qu'il aurait employés s'il avait été dans la peau du personnage; le paysan entendait parler avec des expressions du terroir, le badaud de la grande ville avec un langage peu châtié; assis côte à côte dans la salle, ils forgeaient à leur image les paroles que les acteurs articulaient sans voix. »<sup>211</sup>

Il s'agit là d'un « discours intérieur » au sens littéral du terme : le spectateur prête mentalement une voix au personnage. En optant pour un registre de langue particulier, pour une certaine intonation, etc., le bonimenteur tisse par contre un réseau de connotations qui influencent la façon dont le public se représente les locuteurs diégétiques. L'activité linguistique des spectateurs ne se limitait néanmoins pas à l'attribution de voix, mais s'appliquait aussi à la compréhension générale du message visuel, selon la conception eikhenbaumienne d'un dépassement du signifiant phonétique mental grâce au discours intérieur. C'est pourquoi la comparaison entre discours intérieur et médiation bonimentorielle révèle combien cette dernière exerce une fonction restrictive sur la lecture du film, notamment lorsque l'énonciateur *live* convoie des informations en se basant sur ses propres connaissances paratextuelles ou encyclopédiques. Ce type de savoir intervient également dans l'esprit du spectateur, mais sans la préparation qu'est susceptible d'effectuer le bonimenteur averti qui accompagne pour la nième fois la projection de la même série de bandes.

L'hypothèse d'Eikhenbaum, centrale au sein de son édifice théorique, a plus généralement trait au rôle joué par le langage verbal dans la perception de l'image. La portée de son approche a été notée par Emilio Garroni qui, dans son essai de sémiotique du début des années 70, considérait le texte du Soviétique comme

211. J. A. Keim, *op. cit.*, p. 43.

l'unique précédent notable d'une réflexion relative à l'influence exercée par le langage verbal sur la lecture du film<sup>212</sup>. La question des interactions entre la parole et les images ne concerne cependant pas spécifiquement le boniment, qui n'est à cet égard pas très différent de la voix-over.

Il faut néanmoins noter qu'en passant du «cinéma» parlé au cinéma parlant, la communication verbale directe, effectuée «corps à corps» entre le bonimenteur et les spectateurs, est remplacée par un autre type de relation. Comme le fait remarquer Charles Wolfe, la voix synchrone instituée par le parlant ne constitue pas uniquement un ajout, mais aussi une *soustraction*, celle du corps vivant de l'auteur de la performance<sup>213</sup>. En ce sens, on peut affirmer que, comparativement au cinéma parlé, les voix du parlant sont fondamentalement *over*. Or le procédé spécifique de la voix-over représente le cas le plus ostensible de cette éviction de la dimension physique. C'est en cela que, paradoxalement, la voix-over se rapproche du boniment : sa profération n'étant intrinsèquement pas (illusoirement) rapportée à la visualisation d'un locuteur, elle se pose ouvertement comme un facteur de *médiation* entre le public et la représentation visuelle. On peut dire en effet que toutes les fonctions dont le boniment est susceptible de se charger dans son rapport à l'image valent également pour la voix-over, à l'exception des aspects déterminés par le contexte historique de l'institution du «cinéma» ou par le dédoublement de la représentation visuelle (habillement, mimiques, etc.). Bien que la source de la voix-over ne soit pas ancrée dans l'espace de la salle comme celle du boniment, elle ne provient pas non plus de l'univers diégétique visualisé. Le film peut certes postuler que la voix émane d'un personnage diégétique, mais celui-ci n'a d'existence, à l'instant de la profération, qu'en tant que voix : son insertion dans l'univers du film n'est pas *immédiate* pour le spectateur, mais rappelle le mélange d'extériorité (physique, spatiale) et d'intégration qui caractérise les rapports du bonimenteur au monde de l'écran.

Cependant, il arrive qu'aucun corps ne soit assignable à la voix-over, alors que le bonimenteur, lui, est forcément connu du public, même s'il est provisoirement caché. Le spectateur d'un film parlant à voix-over, fantasmant selon des modalités analysées par Mary

212. Emilio Garroni, *Progetto di semiotica*, Bari : Éditions Laterza, 1972, p. 368.

213. Charles Wolfe, «Vitaphone Shorts and *The Jazz Singer*», *Wide Angle*, Vol. 12, N° 3, juin 1990, p. 64.

Ann Doane l'unité vocale et corporelle d'un «individu», effectue un déplacement: pour reprendre l'heureuse formule de Doane, on peut dire que, lorsque la voix est *over*, «le corps *dans* le film devient le corps *du* film»<sup>214</sup>. La voix-over donne lieu à un phénomène d'*intégration*: la relation au spectateur se voit reversée à l'intérieur même de la représentation. La médiation subsiste, mais elle est naturalisée. La présence corporelle du bonimenteur concurrençait l'image écranique; la voix-over affirme la présence de l'humain sans toutefois constituer un obstacle à la suprématie du monde de l'écran.

En raison des rapports fixes qu'entretiennent voix-over et image, une plus forte détermination réciproque s'instaure entre les composantes sonores et visuelles, et ce d'autant plus lorsque l'une et l'autre ont été créées pour le film, pensées dès leur origine pour être combinées. Il arrive que des films à voix-over comme ceux qui se composent d'images d'archives présentent sur ce point des parentés avec le cinéma bonimenté, parce qu'ils ont été conçus à partir d'images préalables que le verbal vient commenter dans un second temps. Cependant, indépendamment de cette genèse du film – où, de toute manière, montage et récitation verbale sont élaborés de concert, le produit final résulte de liens stricts et immuables entre texte et image. Alors que le bonimenteur s'évertuait, lors d'une performance labile, à commenter une image qui préexistait à son discours, la voix-over et l'image appartiennent à la même unité textuelle dont elles composent des strates d'énonciation distinctes, mais foncièrement complémentaires. Le fait que la voix-over s'origine, contrairement au boniment, dans le pôle machinique du dispositif cinématographique est à la fois le garant de son intégration à la représentation visuelle – qui est également le produit d'un défilement mécanique – et la cause d'un changement important au niveau de la vocalité. Le caractère réitérable de l'enregistrement de la voix abolit la «tactilité» de ce phénomène éphémère qu'est la transmission orale pour Paul Zumthor. Il ne faut cependant pas dénier à la voix-over toute dimension performantielle: l'acteur qui prête sa voix pour narrer un film *over* manifeste à des degrés variables sa «présence» par les caractéristiques de sa voix. Cette vocalité renferme toujours une part de la corporéité

214. Mary Ann Doane, «The Voice in the Cinema: The Articulation of Body and Space», *Yale French Studies*, N° 60, 1980, p. 35.

du locuteur, fût-il absent à la fois de la salle et de l'image du film. Car la voix-over, contrairement aux voix synchrones qui sont plus directement asservies à l'action visualisée, conserve une certaine autonomie qui lui permet d'être finement travaillée en tant que phénomène vocal. Il n'en reste pas moins que l'enregistrement, en dépit de son indéniable valeur indiciare, résulte d'une lecture effectuée dans l'isolement d'un auditorium, alors que le boniment, dont je propose d'étudier dès à présent les diverses fonctions, renvoie à une véritable performance.

### II.3 LES FONCTIONS DU BONIMENT

Afin de mieux saisir le rôle joué par la voix lors des projections animées des premiers temps, il me semble nécessaire d'établir et de discuter une typologie des fonctions bonimentorielles à l'intérieur du cadre général du dispositif audiovisuel. Dans le cinquième chapitre de son ouvrage sur la question («Les fonctions et la performance du bonimenteur»), Germain Lacasse annonce une même entreprise. Toutefois, il accorde peu de place aux fonctions proprement dites qu'il décrit, dans une volonté d'établir une périodisation dont on peut à mon avis faire abstraction pour atteindre plus de généralité, comme appartenant à quatre grands types (boniment de vente, d'explication, de traduction, boniment performatif), et non comme le résultat d'une combinaison de paramètres – ainsi la «performativité» vaut-elle par exemple tout autant pour le premier type. Je propose quant à moi une approche d'obédience linguistique qui permet de distinguer deux niveaux d'analyse: celui de la communication audiovisuelle dans son ensemble et celui, plus restreint, des relations entre paroles et images. Comme cette seconde question n'est pas spécifique au «cinéma» parlé, je ne l'aborderai qu'ultérieurement dans une optique «textuelle» à propos du cinéma parlant. Par contre, je développerai ici la notion de «traduction» proposée par Germain Lacasse, car elle permet d'articuler les relations verticales internes à la représentation sur les relations horizontales avec l'audiospectateur.

Dans un article récent, Jean-Pierre Sirois-Trahan convoque, entre autres modèles, celui de la communication verbale de Roman Jakobson pour montrer que, dans la diversité des fonctions considérées par le linguiste, seule la fonction référentielle prédomine au sein de la «théorie du dispositif» développée dans le

champ du cinéma. Cette réduction résulterait du fait que les théoriciens s'appliquent presque exclusivement à décrire le spectateur d'une « diégèse narrative »<sup>215</sup>. Le « dispositif » décrit par Baudry ou Metz en termes de régime psychique du spectateur ne correspond pas au sens que je donne à cette notion dans le présent ouvrage<sup>216</sup>. L'approche psychanalytique ne s'attache tout au plus qu'à l'une des facettes, celle qui est centrée sur le pôle spectatoriel. Il me paraît par contre intéressant d'exploiter plus avant les notions théorisées par Jakobson qui, bien que se donnant pour objet la conceptualisation linguistique de la poésie, offre un cadre général qui prend en compte le verbal dans une situation de communication<sup>217</sup>. En effet, Jakobson définit six facteurs qu'il estime définitoires pour toute communication verbale, et auxquels il associe une fonction langagière spécifique: le destinataire (fonction émotive), le destinataire (fonction conative), le message (fonction poétique), le contexte (fonction référentielle), le contact (fonction phatique) et le code (fonction métalinguistique). Chaque fonction, présente à des degrés variables dans tout énoncé, est donc définie par rapport à celui des six pôles sur lequel l'utilisateur du langage décide de mettre l'accent. L'avantage pour une optique combinatoire comme la nôtre – nécessaire au vu de la diversité des pratiques bonimentorielles et de l'impossibilité de statuer aujourd'hui avec exactitude sur ce qu'elles furent – réside dans le fait que ces catégories ne s'excluent pas mutuellement, mais entretiennent des liens hiérarchiques. Cet aspect est particulièrement marqué en ce qui concerne la distinction entre phatique et conatif: le « contact » implique intimement le destinataire, si bien que la première de ces fonctions s'accompagne généralement de la seconde. On pourrait établir dans de tels cas que le *phatique* se mue en conatif dès lors que l'énoncé est suffisamment long pour que son rôle excède celui qui consiste à s'assurer de l'attention de l'auditeur. Plus généralement, toute fonction qui est prédominante à un moment donné du discours verbal peut être ensuite reléguée au second plan, tout en continuant partiellement à se manifester. La proportion relative

215. Jean-Pierre Sirois-Trahan, « La bataille du parlant, rue Sainte-Catherine, vingt ans avant le *Jazz Singer* », in A. Gaudreault (dir.), *Au pays des ennemis du cinéma...*, op. cit., pp. 157-158.

216. Jean-Louis Baudry, « Le dispositif: approches métapsychologiques de l'impression de réalité », *Communications*, N° 23, 1975; Christian Metz, *Le Signifiant imaginaire*, Paris: Christian Bourgois, 1993 [1<sup>re</sup> édition 1977].

217. Roman Jakobson, « Linguistique et poésie », in *Essais de linguistique générale*, Paris: Minuit, 1963 [1<sup>re</sup> édition 1960].

de chaque fonction permet ensuite d'esquisser le portrait d'une actualisation particulière du langage au sein d'une situation d'oralité.

La fonction *émotive* du boniment se caractérise par la présence d'éléments d'expressivité qui ne se limitent pas à l'affirmation du «je»: Jakobson mentionne autant les niveaux phonique et grammatical que lexical. Il s'agit en fait du «ton» pour lequel opte le bonimenteur, par exemple l'emphase ou l'ironie. Cette «coloration» du propos résulte directement de l'idiosyncrasie du locuteur et, souvent, de caractéristiques vocales (accent, intonation, débit, etc.); elle est donc fortement ancrée dans le présent du spectacle. Jakobson précise par ailleurs que l'expression transmise peut être aussi bien vraie que feinte<sup>218</sup>. Cette posture du sujet parlant constitue un facteur important de médiation entre l'audiospectateur et la représentation visuelle. On peut imaginer que, selon la nature du montré, le bonimenteur se fasse plus ou moins discret. L'identité du bonimenteur et d'un «personnage» à l'écran devait sur ce point constituer une situation particulière: la médiation spectateur↔représentation pouvait alors reposer sur la combinaison de deux fonctions prédominantes (émotive et référentielle). Comme certains historiens l'ont relevé, il arrivait que les projections alternassent avec des performances *live* d'acteurs apparaissant également dans le film, comme dans le «*rensageki*» japonais<sup>219</sup> ou dans les spectacles de Cor Schuring à Amsterdam<sup>220</sup>; toutefois, le bonimenteur occupe dans de tels cas une position d'extériorité par rapport au spectacle, alors que sa subjectivité est intimement liée à ce qui est montré lorsque, dans des films comme les «*travelogues*», il incarne de surcroît l'instance de réalisation de la bande. Charles Musser, qui a étudié ce «genre», affirme que «l'identification du public» aux conférenciers de *travelogues* s'opère à trois niveaux: le *showman* est à la fois celui que l'on voit sur l'écran, celui qui tient l'appareil de prise de vues et celui qui parle sur la tribune<sup>221</sup>. Ces trois modes de présence ne sont toutefois pas du même ordre, et n'impliquent pas forcément une «identification», du moins pas

218. *Ibid.*, p. 214.

219. Cf. Hiroshi Komatsu et Frances Loden, «Mastering the Mute Image: The Role of the *Benshi* in Japanese Cinema», *Iris*, N° 22, 1996, p. 41.

220. Cf. Ivo Blom et Ine van Dooren, «Ladies and Gentlemen, Hats off, Please!», *Iris*, N° 22, 1996, p. 87.

221. Charles Musser, «The Travel Genre in 1903-1904. Moving towards Fictional Narratives», in Thomas Elsaesser (dir.), *Early Cinema. Space, Frame, Narrative*, London: BFI, 1990, p. 127.

dans le sens que l'on donnera à ce terme pour aborder un état ultérieur du cinéma. Ce qui pourrait constituer une identification primaire et secondaire au sens de Metz<sup>222</sup> est en fait subordonné à la présence du bonimenteur qui, par son apparence physique et ses dires, offre le point de repère nécessaire à la reconnaissance des deux autres instances. Le spectateur *l'identifie* dans l'image, ou comme étant la source de celle-ci, mais ne *s'identifie* pas forcément à lui. Il en allait différemment des « ciné-déclamations » russes que Valérie Pozner décrit ainsi : « Ce spectacle consistait à doubler en direct des films spécialement prévus à cet effet : l'acteur que l'on voyait jouer était généralement celui-là même dont on entendait la voix dans la salle. »<sup>223</sup> Dans ce cas, le synchronisme vocalabial assurait l'inscription de l'instance verbale dans la diégèse filmique – il n'y a plus ni tribune ni présence du projecteur – et renforçait l'effet de réel pour les audiospectateurs auxquels la voix de l'acteur de l'écran était familière. Par contre, la médiation bonimentorielle des *travelogues* pouvait recourir de façon optimale à la fonction émotive, puisque la présence à l'image du conférencier légitimait la dimension subjective de la transmission du savoir. Comme il « a été là », le bonimenteur peut ensuite, dans le présent de sa performance, parler des vues comme de son vécu. En outre, l'association du locuteur avec le filmage permet d'humaniser la dimension mécanique. Bien sûr, les *travelogues* offrent une situation singulière, puisque la majorité des autres bandes que les exploitants achetaient ou louaient ne présentaient pas cette possibilité d'un lien aussi direct entre le bonimenteur et la représentation. Deux orientations se présentent néanmoins au bonimenteur : soit il soumet la fonction référentielle à la fonction émotive, ce qui permet de prendre une certaine distance face aux images ; soit il opte pour la posture inverse, et insère de la subjectivité à l'intérieur de la diégèse. Ce second cas trouve son expression la plus évidente dans la reproduction de dialogues, qui peut conduire le locuteur à contrefaire sa propre voix pour l'adapter aux personnages de l'écran. Cette pratique était courante au Japon, comme l'indique la *benshi* Sawato Midori dans un entretien<sup>224</sup>.

Le pôle opposé qu'est le destinataire est pris en compte dans le modèle de Jakobson à travers la fonction dite « *conative* ». Il s'agit

222. Ch. Metz, *Le Signifiant imaginaire*, op. cit., pp. 61-81.

223. Valérie Pozner, « La ciné-déclamation en Russie », art. cit., p. 136.

224. In Larry Greenberg, *The Benshi – Japanese Silent Film Narrators*, Tokyo: Urban Connections, 2001, pp. 13-14.



de tout ce qui relève de l'adresse, l'un des principaux traits à la fois de l'oralité en général, mais aussi de l'esthétique visuelle du cinéma des premiers temps. On distinguera l'exhibition momentanée de la relation au spectateur (brève prise de contact relevant du phatique) d'un régime spectaculaire (le « cinéma des attractions »), voire d'un dispositif (le « cinéma » parlé) globalement dominé par le mode de l'adresse, et donc par la fonction conative. À l'époque du cinéma des premiers temps, l'interpellation du public devait varier selon le lieu de projection et le rôle assigné au « bonimenteur » : la harangue de l'aboyeur situé à l'extérieur de l'établissement relève par exemple majoritairement de la fonction phatique. Toutefois, les fonctions phatique et conative ne caractérisent pas seulement ce que Lacasse nomme « le boniment de vente » – même s'il est vrai que ce dernier devait posséder une charge informationnelle minimale – mais, plus généralement et dans des proportions variables, toute intervention bonimentorielle. Une chronique de *La Presse* du 27 août 1907 renvoie littéralement à la notion de Jakobson : « Le clou de la représentation était, certainement, *la prise de contact* de l'humoriste A. de Reusse avec l'élégante clientèle du National Biograph. »<sup>225</sup> Au cours du spectacle, la fonction phatique permet au bonimenteur de s'assurer (de) l'attention du public et de renforcer le lien de complicité que le dispositif favorise.

L'importance relative de ces fonctions dépend en partie de l'emplacement temporel du boniment. On distinguera les cas où il se déroule *avant* les projections (primat du phatique), *après* (primat du conatif) ou *pendant* (primat du contexte référentiel). Les « bornes » du film devaient être par contre plus propices à la réflexivité, à l'instar des génériques dans le cinéma ultérieur. Toutefois, la définition jakobsonienne de la fonction *métalinguistique* mérite d'être nuancée dans le cas d'un spectacle audiovisuel : est-ce à son propre discours que se réfère le bonimenteur, ou à certains codes spécifiques à la lecture de l'image ? Il peut s'agir soit d'un métalangage strict dans le sens où l'entend Jakobson (par exemple lorsque le bonimenteur donne une définition d'un terme qu'il vient d'employer), soit d'explications verbales relatives au « signifiant » visuel, notamment lorsqu'il utilise des expressions comme « avant-plan » ou « vue suivante ». Certes, le film devient alors le « référent » de la parole, le bonimenteur exploitant l'aptitude du langage verbal,

225. Cité par G. Lacasse, *Le Bonimenteur...*, *op. cit.*, p. 126.

discutée par Christian Metz, à faire office de métacode, y compris par rapport à des codes non linguistiques comme ceux qui sont à l'œuvre dans le signe iconique<sup>226</sup>. On conviendra toutefois qu'une différence notable existe entre la prise en compte du code linguistique et la corréférence verbo-iconique. C'est pourquoi je propose, afin de ne pas galvauder l'expression de Jakobson pour qui le métalangage implique la seule référence au verbal à travers le langage verbal, le terme plus générique de fonction *autoréférentielle* (ou plus simplement *réflexive*) lorsque le bonimenteur se réfère au « langage » des images. On peut élargir cet objet du discours au dispositif cinématographique lui-même, c'est-à-dire non seulement au signifiant iconique convoqué par la représentation de l'écran, mais aussi à la machinerie, à « l'espace de la production » (qui mène entre autres à l'autodésignation lorsque le bonimenteur parle de sa propre fonction de médiation) et au spectateur (pris à partie non pas en tant qu'être vivant en général, mais en tant que spectateur particulier d'un film donné). Rien n'interdit en effet au bonimenteur de mentionner dans son discours l'appareil de projection, le pianiste ou l'exploitant. Quant à la dimension plus spécifiquement métacodique de la fonction autoréférentielle, elle correspond probablement à l'un des rôles les plus importants que le bonimenteur a dû jouer : celui, comme le dit Châteauevert, d'« éduquer le regard du spectateur »<sup>227</sup> afin qu'il puisse intégrer la nouvelle culture visuelle de l'image en mouvement, puis la lecture des bandes pluri-punctuelles. Dans le contexte historique du cinéma parlé, le gradient métalinguistique me semble plutôt dépendre de la fonction sociale du verbal. Les discours de l'époque reflètent par exemple de telles variations à travers l'opposition entre le « bonimenteur » et le « conférencier ».

La réflexivité vaut théoriquement pour le code comme pour le message, qui, dans le cinéma parlé, est à la fois iconique et verbal. Toutefois, la fonction *poétique* de Jakobson est trop fortement liée à la dimension linguistique de la communication – ou du moins langagière – pour s'appliquer aux films des premiers temps qui travaillaient peu « le côté palpable des signes »<sup>228</sup>. Ce n'est qu'une fois sa légitimité artistique acquise et les spécificités de son langage établies que le médium pourra prétendre faire fonctionner le message pour son propre compte, de façon plus ou moins indépendante de la

226. Ch. Metz, « Le perçu et le nommé », *art. cit.*, pp. 147-150.

227. Jean Châteauevert, « Le cinéma muet », *Iris*, N° 22, 1996, p. 107.

228. R. Jakobson, « Linguistique et poétique », *art. cit.*, p. 218.

nature du représenté. Au début du siècle importe avant tout le contenu des vues, rarement la façon dont elles sont montrées; or l'évolution du cinéma vers un « langage » propre qui s'effectue à travers les discours de légitimation appartient à une période postérieure au déclin du bonimenteur. Dans les projections bonimentées, la fonction poétique était donc prioritairement endossée par la parole, le bonimenteur pouvant travailler ses phrases au niveau de leur rythme, de la sonorité des mots, etc. Jakobson précise bien que la fonction poétique ne se réduit pas au genre dans lequel on lui accorde une fonction dominante, c'est-à-dire la poésie: on peut imaginer que le plaisir de l'écoute de certains bonimenteurs provenait d'un accent ponctuellement placé sur cette fonction poétique, lieu d'une forte exploitation du potentiel de la vocalité. Assonances, allitérations, paronomases, etc. acquièrent une dimension particulière lorsque le texte est proféré, l'oralité actualisant véritablement la « courbe mélodique » que peut suggérer l'écrit (c'est pourquoi Zumthor parle globalement de « poésie orale »). Ainsi n'est-ce pas un hasard si Roman Jakobson, qui a développé sa conception de la poésie dans un cadre communicationnel, souligne une vingtaine d'années plus tard la nécessité de distinguer poésie écrite et récitée dans sa préface à l'ouvrage *La Vive voix* de Fónagy<sup>229</sup>.

La dernière fonction proposée dans les *Essais de linguistique générale* apparaît comme la plus évidente, puisque l'on dit ordinairement que le bonimenteur « commente » les vues: la fonction *référentielle* sert à compléter, à reformuler (*via* une trans-sémiotisation) et à clarifier ce que montre l'image. Le référent verbal peut soit déborder le contenu des images, soit sélectionner certaines informations qu'elles présentent, soit développer un discours qui n'entretient aucun lien direct avec ce qui est montré. La thèse centrale selon laquelle le boniment aurait contribué à linéariser les films postule le rôle prépondérant de la fonction référentielle, rapportée essentiellement à l'édification d'une diégèse et au renforcement de la continuité narrative. On touche alors aux procédures régissant la conjonction du verbe et de l'image, qui ne diffèrent pas fondamentalement dans la projection bonimentée de ce que l'on trouvera dans les films parlants sous la forme de la voix-over (voir VI. 5).

229. Ivan Fónagy, *La Vive voix*, Paris: Seuil, 1983, p. 8.

Si l'on veut poursuivre, sur le plan sémantique et non plus rythmique, l'analyse de l'axe bonimenteur-représentation, il est nécessaire d'ajouter aux fonctions générales empruntées à la linguistique de Jakobson des sous-fonctions permettant de tenir plus spécifiquement compte de l'influence du verbal, celui-ci engageant des principes qui déterminent la sélection, la hiérarchisation, la mise en relation ou l'adaptation d'informations visuelles. L'approche de ces phénomènes gagne selon moi à exploiter des acquis théoriques développés dans un autre champ d'étude, celui de la traduction (littéraire). Certes, Germain Lacasse se refuse à considérer le bonimenteur comme un traducteur, opposant à ce terme celui, tout à fait en accord avec l'optique défendue ici, de « médiateur ». Toutefois, lorsqu'il affirme que « le mot traducteur conviendrait, à la condition de reconnaître que traduire est impossible et que le traducteur réécrit le texte ou le recompose dans sa propre langue », Lacasse témoigne d'une conception restreinte de la « traduction »<sup>230</sup>. Car la traduction littéraire se présente, plus que tout autre type traductionnel, comme un processus d'interprétation et de recreation, comme l'est par ailleurs, dans une certaine mesure, toute traduction (ce que semble négliger Lacasse). Dans sa classification des textes d'accompagnement, l'auteur du *Bonimenteur des vues animées* mentionne par contre le « boniment de traduction »<sup>231</sup> : il s'agit de la situation particulière où, précisément, le bonimenteur traduit le texte des intertitres dans l'idiome des spectateurs. Notons que, même dans de tels cas où n'intervient aucune transsémiotisation, la traduction n'est pas forcément littérale. De plus, lors du transfert d'une langue-source vers une langue-cible, certaines modifications sont inévitables au niveau du registre de langue, du lexique, de l'accentuation des mots, etc. Le passage de l'écrit à l'oral n'est pas sans conséquences, à moins que les intertitres (par exemple de dialogues) ne présentent déjà des effets d'oralité. Cette tâche ne constituait pas l'exclusivité de la fonction du bonimenteur, mais intervenait parmi diverses autres sous-fonctions, sans quoi l'improvisation du bonimenteur s'en serait trouvée fortement limitée. Les historiens ont fréquemment relevé que les exploitants devaient adapter la projection à un public illettré ou d'immigrants ne maîtrisant pas la langue nationale, ou simplement faire comprendre des films

230. G. Lacasse, *Le Bonimenteur...*, op. cit., p. 131.

231. *Ibid.*, p. 128.

étrangers pour lesquels les cartons n'avaient pas été traduits dans la langue des spectateurs. Ce besoin tout pratique ouvre la voie à une réappropriation du texte écrit – mais aussi du texte filmique – par le bonimenteur. Anse van Beusekom a par exemple constaté pour les Pays-Bas, en dépouillant les offres d'emploi pour bonimenteurs, qu'en dépit du besoin de traduire les intertitres de films provenant essentiellement de France, aucune connaissance de la langue française n'était requise<sup>232</sup>. Le glissement de la traduction des intertitres au sens strict vers un processus plus large qui repose sur une lecture de l'image devait être fréquent. La « traduction » effectuée à travers le boniment implique bien sûr un transfert de l'information convoyée visuellement vers un autre système de signes: il s'agit, pour reprendre la terminologie de Jakobson, d'un cas de « traduction intersémiotique ou transmutation »<sup>233</sup>. Les changements impliqués par l'activité traductionnelle sont donc nécessairement plus importants qu'en cas d'identité de langue (reformulation) ou de nature sémiotique (traduction interlinguale).

Les critères qui régissent les processus de traduction ne sont pas uniquement linguistiques, mais relèvent également de la pragmatique. Blom et Van Dooren rapprochent la fonction bonimentoriel de cette acception plus large de la traduction lorsqu'ils disent des spectateurs du cinéma parlé qu'ils « demandaient des adaptations – une « traduction » dans leur propre culture, leur propre langue ou même leur propre dialecte »<sup>234</sup>. En effet, au vu des déterminations réciproques qui s'établissent entre configurations langagières et représentations du monde, le passage d'une langue à l'autre implique inévitablement un « transfert culturel » qui tient compte des modifications nécessitées par le nouveau contexte (sociologique, géographique, etc.) qui l'accueille. Ce phénomène, patent dans le cas de sociolectes (régionalisme, jargon, dialecte), intervient en fait à tous les niveaux de la traduction, à proportion de l'écart existant entre la culture-source et la culture-cible, dont dépend l'importance des ajouts ou des transformations. Dans un bref essai, j'ai tenté de classer et de discuter un certain nombre de paramètres applicables à l'analyse de la traduction littéraire, que je

232. A. van Beusekom, « The Rise... », *art. cit.*, p. 132.

233. Roman Jakobson, « Aspects linguistiques de la traduction » [1959], in *Essais de linguistique générale*, Paris: Minuit, 1963, p. 79.

234. I. Blom et I. van Dooren, « Ladies and Gentlemen... », *art. cit.*, p. 89.

propose de reprendre ici<sup>235</sup>. Il m'est apparu que les variations culturelles représentent un facteur qui conditionne de nombreux choix de traduction.

Les diverses solutions de traduction ressortissent à quatre opérations principales qui sont également à l'œuvre dans le boniment. Comme le traducteur, le bonimenteur peut *reprendre* « littéralement » le « texte » original, *supprimer* (ou plutôt *occulter*) certaines informations jugées superflues, notamment lorsque le film se rapporte à la culture-cible et donne des précisions qui apparaissent désormais comme des lapalissades pour l'auditeur, effectuer une *transposition* visant à expliciter certains éléments qui n'auraient pas la même évidence pour les nouveaux destinataires, ou offrir un *supplément* d'informations, soit sur un mode naturalisé où l'explication est intégrée à la traduction, soit de façon plus manifeste en interrompant le récit par un commentaire explicatif. La fonction référentielle est alors intimement liée à la fonction phatique, le bonimenteur ne pouvant se permettre de perdre le contact avec ses auditeurs. Ces choix sont par contre passablement subjectifs, et dépendent de la façon dont le « traducteur » se représente sa tâche et son public. Il se peut par exemple que la visée de l'ajout explicatif dépasse les seules nécessités sémantiques du pont interculturel ou du contexte narratif pour s'inscrire dans une démarche pédagogique. Tout film pouvait faire office de support à la transmission d'un savoir sur le monde, le bonimenteur déployant des inférences « encyclopédiques »<sup>236</sup> à partir de l'un ou l'autre élément visualisé, indépendamment de la pertinence strictement narrative de ces informations. Malgorzata Hendrykowska mentionne l'exemple du Polonais Libanski, qui commentait les technologies représentées dans les fantaisies de Méliès à des fins de vulgarisation scientifique<sup>237</sup>. Explication et narration constituent les deux modes principaux d'intégration des vues isolées par le pouvoir homogénéisant du continuum temporel de la performance.

À ces aspects d'ordre pragmatique s'ajoutent deux autres niveaux auxquels opèrent également les procédés traductionnels : celui du signifiant et celui du signifié. Il va de soi que le premier ne

235. Alain Boillat, *Deux traductions de « La Chevelure de Bérénice » de Claude Simon. Paramètres pour l'analyse de traductions littéraires*, Lausanne : Centre de Traduction littéraire, 1999.

236. Au sens d'Umberto Eco, *Lector in fabula. Le Rôle du lecteur*, Paris : Grasset, 1985 [1<sup>re</sup> édition 1979], p. 33.

237. Malgorzata Hendrykowska, « Professeur, acteur ou bouffon ? Le bonimenteur sur les territoires polonais avant 1914 », *Iris*, N° 22, 1996, p. 167.

doit pas être considéré, dans le cas du boniment, d'un point de vue linguistique (à moins de ne se concentrer que sur la traduction des intertitres) : la question de la « construction syntaxique » et du « rythme » – critères listés dans la catégorie du « signifiant »<sup>238</sup> – subsiste néanmoins dans les limites imposées par la différence sémiotique entre verbal et image. À l'organisation « syntaxique » du texte scriptural correspond au cinéma la mise en chaîne des « plans » qui, dans le cinéma des premiers temps, pouvait résulter d'une juxtaposition, effectuée par l'exploitant, de bandes issues de contextes différents. On peut alors imaginer que le bonimenteur usait de la fonction (auto)référentielle de son intervention pour rendre compte des articulations opérées au niveau de l'image. On constate alors – contrairement à la traduction linguistique où les deux signifiants sont de même nature – un changement de niveau, où le signifiant filmique devient le signifié du verbal.

Si la comparaison des signifiants montre rapidement ses limites lorsqu'il s'agit d'une traduction intersémiotique, il n'en va pas de même des signifiés, indépendants du support de la communication. De la nature polysémique de l'image ciné-photographique découle l'importance du rôle sémantique du verbal, qui permet de sélectionner certains signifiés iconiques parmi ceux que présente l'image. Le boniment endosse alors la fonction « d'ancrage » dont Roland Barthes a parlé à propos des relations entre messages linguistique et iconique, précisant que « le texte *dirige* le lecteur entre les signifiés de l'image, lui en fait éviter certains et en recevoir d'autres »<sup>239</sup>. À ce niveau de la dénotation s'ajoute celui de la connotation : le verbal est particulièrement apte à convoquer un sens second qui permet de conférer au discours une valeur supplémentaire d'ordre affectif, axiologique, etc. Le sens connoté permet ainsi au bonimenteur d'instaurer une véritable proximité avec son public en créant des phénomènes d'empathie avec ce qui est montré, ou en adoptant le même esprit de communauté « idéologique » que les destinataires. L'hypothèse centrale de l'ouvrage de Germain Lacasse sur le bonimenteur, selon laquelle le commentateur contribuait à retourner le « message » du film contre ceux-là mêmes qui l'avaient conçu (c'est-à-dire contre « l'institution »), implique de la part de la médiation ce type de « modulation sémantique ».

238. A. Boillat, *Deux traductions...*, *op. cit.*, p. 57.

239. R. Barthes, « Rhétorique de l'image », *art. cit.*, p. 44.

L'ajout de connotations n'enraye pas nécessairement le récit – et donc la fonction linéarisante du boniment – car une structure narrative peut tout à fait s'accommoder de jugements de valeur, notamment dans la phase de résolution. Lorsque, selon Jean Châteauvert, «le conférencier tire les enseignements moraux du film»<sup>240</sup>, son activité participe de la fonction configurationnelle que Jean-Michel Adam conçoit, à la suite de Paul Ricoeur, comme étant la dernière séquence narrative d'un récit, celle de l'«évaluation finale ou morale»<sup>241</sup>. Tout comme les fonctions de Jakobson, les divers types d'organisation textuelle (modes narratif, informatif, explicatif, descriptif, etc.) doivent être pensés en termes de «dominantes», non comme des catégories mutuellement exclusives. En procédant majoritairement à des ajouts utiles à la lisibilité et à l'intelligibilité du film, le bonimenteur devait passer fréquemment d'un mode à l'autre.

Au niveau sémantique, un autre critère paraît important dans le cadre de l'analyse de la traduction littéraire qui tient à l'appartenance du texte à un type de discours<sup>242</sup>. Or Germain Lacasse souligne cet aspect intertextuel et pragmatique à propos du bonimenteur, dont la «conférence imposait une certaine consigne de lecture et certaines attentes qui sont généralement le fruit de la notion de genre»<sup>243</sup>. Un tel usage du langage fortement phatique constitue nécessairement l'un des faisceaux de déterminations qui agit sur le «contrat de lecture» passé avec le spectateur, et ceci d'autant plus que le bonimenteur, en accompagnant les films des premiers temps qui ne présentaient pas la «clôture» des organisations narratives ultérieures, était susceptible d'orienter fondamentalement la réception en optant pour l'une des diverses lectures autorisées par le film. Ces choix impliquent la définition d'un horizon d'attente, d'un mode d'organisation, voire d'un régime de croyance spécifique. Charles Musser a souligné, à propos de films de voyage réalisés par Porter pour la compagnie Edison, qu'ils pouvaient être regardés comme des documentaires sur un lieu (par exemple Coney Island dans *Boarding School Girls*, 1905) ou comme des films narratifs qui mettent en scène des personnages se déplaçant

240. J. Châteauvert, «Le cinéma muet», *art. cit.*, p. 107.

241. J.-M. Adam, *Le Récit*, Paris: PUF, p. 93.

242. A. Boillat, *Deux traductions...*, *op. cit.*, p. 58.

243. Germain Lacasse, «Palimpsestes, trame et traces: tradition orale et notion de genre dans le cinéma des premiers temps», *Iris*, N° 20, 1995, p. 69.



dans ce décor ; il incombait au bonimenteur de mettre l'accent sur l'une des deux dimensions<sup>244</sup>. Ainsi, la mise en place d'oppositions aussi importantes que narratif *vs.* non narratif ou documentaire *vs.* fiction pouvait être le fait de la posture choisie par le bonimenteur. La double perspective évoquée par Musser correspond en fait à la dichotomie monstration/narration : soit le bonimenteur exploite le potentiel narratif de certaines vues et fait référence à des personnages récurrents, soit il véhicule des informations à propos de la nature du montré. Ces deux modes peuvent être « programmés » par le film (comme on le constate dans certains descriptifs des maisons de production), mais ils ne sont actualisés qu'avec le boniment (et laissés au choix du spectateur dans le cas de projections qui en sont dépourvues). Le boniment oscille alors entre le spectacle pur et l'explication, puisque son auteur peut manifester un enthousiasme envers des éléments isolés dont il fait l'éloge (les « attractions » de Coney Island ou du Luna Park se prêtant particulièrement bien à ce type de rapport au film) ou inscrire les différents « monuments » ou événements montrés dans une structure verbale qui assure des liens plus ou moins solides entre eux. La non-fiction n'est donc pas moins adaptée aux phénomènes d'intégration que la fiction. Stephen Bottomore a d'ailleurs montré que le film d'actualité fut, dès les premières années du cinématographe, un terrain privilégié pour le développement d'un « montage » en termes de juxtaposition de vues relatives à un même événement<sup>245</sup>. Le boniment joua vraisemblablement un rôle considérable au niveau de la mise en relation des vues.

L'analyse fonctionnelle du boniment permet d'évaluer la richesse de l'apport fourni par l'accompagnement verbal. Face à l'absence de traces concrètes de cette pratique orale, le large spectre couvert par le modèle communicationnel de Jakobson ainsi que la méthodologie issue de la théorie de la traduction permettent d'en imaginer les orientations possibles. Néanmoins, cette approche verbocentriste ne rend pas compte des particularités « machiniques » du dispositif cinématographique. C'est cet aspect de la médiation bonimentorielle qui retiendra mon attention dans la section suivante.

244. Ch. Musser, « The Travel Genre... », *art. cit.*, p. 125.

245. Stephen Bottomore, « Shots in the Dark. The Real Origin of Film Editing », in Th. Elsaesser (dir.), *Early Cinema...*, *op. cit.*

## II. 4 LA MEDIATION SPECTATEUR ↔ MACHINERIE : L'« IDEOLOGIE » DES DISCOURS SUR LE BRUIT DU PROJECTEUR

La manifestation la plus patente de la relation entre le boniment et la machinerie réside dans la présentation verbale de l'appareil de projection, une pratique que l'on trouvait déjà chez Kircher. L'exploitant exhibe avant tout l'innovation techn(olog)ique, les images projetées important plus en elles-mêmes que pour leur contenu. Joseph Anderson et Donald Richie ont établi que les projections au Japon des (tout) premiers temps, dans la filiation des conventions théâtrales du kabuki (où les changements de décor s'effectuent avec les rideaux levés), consistaient fréquemment à exposer la mécanique du dispositif : comme l'écran se trouvait sur la gauche de la « scène » et le projecteur sur la droite, l'intérêt visuel de ce dernier était mis sur un pied d'égalité avec celui de l'image projetée<sup>246</sup>. Noël Burch, toujours prompt à souligner la possibilité d'une rupture avec la tradition représentative occidentale qui domine dès le milieu des années 10, corréle ce mode de présentation avec l'art du *benshi*, et note que l'exhibition du dispositif de projection et du « présentateur-conférencier » avait également cours en Occident<sup>247</sup>. Selon cette théorie, le boniment devient le parangon d'une démarche anti-illusionniste qui se serait perdue en Occident, alors qu'elle aurait durablement marqué l'exploitation cinématographique nipponne.

Si ce type de médiation où le bonimenteur fait porter son discours sur la machinerie a pu réapparaître sporadiquement au gré des développements technologiques ou d'assertions autoréférentielles ponctuelles, il ne semble avoir tout au plus prédominé que durant les deux premières années du cinématographe, phase de l'évolution du médium qui correspond à la « *novelty period* » discutée par Tom Gunning<sup>248</sup>. Charles Musser affirme quant à lui que « la période de nouveauté du cinéma n'a duré qu'un an, de la fin 1895 au début 1897, soit l'équivalent d'une saison de programmation »<sup>249</sup>. Alors qu'à ses tout débuts, l'appareil de projection de

246. Joseph L. Anderson et Donald Richie, *The Japanese Film: Art and Industry*, Princeton: Princeton University Press, 1960, pp. 23-24.

247. Noël Burch, *Pour un observateur lointain*, Paris: Gallimard; Éditions des Cahiers du cinéma, 1982 [1<sup>re</sup> édition 1979], p. 85.

248. Tom Gunning, « The Scene of Speaking... », *art. cit.*, p. 73.

249. Charles Musser, « Pour une nouvelle approche du cinéma des premiers temps: le cinéma d'attractions et la narrativité », in Jean Gili, Michèle Lagny, Michel Marie et Vincent Pinel (dir.), *Les Vingt Premières Années du cinéma français*, Paris: AFRHC, 1995, p. 162.

vues animées constitue en lui-même une attraction, le spectacle cinématographique évolue assez rapidement, ainsi que le note Lacasse, «vers une intégration de l'attraction à l'univers diégétique»<sup>250</sup>, ce qui entraîne nécessairement, en ce qui concerne la fonction du boniment, le passage du métafilmique au référentiel. L'intérêt de Noël Burch pour un dispositif dont la nature «exhibitionniste» repose sur la voix du bonimenteur révèle l'importance théorique que peut revêtir la question de la médiation bonimentorielle dans son rapport à la dimension machinique du spectacle cinématographique. Dans son étude ultérieure sur le cinéma des premiers temps, Burch discute le rôle de la musique d'accompagnement – dans ce contexte passablement comparable à celui du boniment – en tant que facteur de «mise en ordre» du déroulement des séances, notamment parce qu'elle «sert à isoler le spectateur du bruit du projecteur, des toux, des commentaires chuchotés, etc.»<sup>251</sup>. Contrairement aux autres perturbations sonores issues du pôle (et de l'espace) spectatorial, le bruit provoqué par l'appareil de projection a des implications qui touchent à l'ensemble des paramètres définissant le dispositif. Il n'est dès lors pas étonnant qu'il soit pris en compte de façon récurrente par les théoriciens du son au cinéma.

Les incidences de la médiation bonimentorielle sur la perception de la machinerie ne sont pas uniquement liées à la fonction autoréférentielle du verbal. Il me semble en effet que la seule présence humaine du bonimenteur questionne et transforme le rapport du spectateur à la machinerie. Si l'on prend vraiment la mesure des spécificités «unificatrices» de la transmission orale telle que l'envisage Paul Zumthor, on peut saisir comparativement en quoi réside l'«étrangeté» de la projection animée muette. La mécanique constitue en effet une figure d'altérité dans une représentation visuelle majoritairement centrée sur l'humain (et cela d'autant plus lorsque les films se narrativisent) et au sein d'un spectacle considéré comme *vivant*. Cette altérité, la voix vive est incapable, en dépit de la dynamique qu'elle engendre avec son auditoire, de l'annihiler totalement, mais elle peut la nier ou l'occulter. Pour le spectateur du cinéma des premiers temps, les productions sonores *live* fournissent, tout autant (et peut-être plus) qu'un apport

250. Germain Lacasse, «Du cinéma oral au spectateur muet», *Cinémas*, Vol. 9, N° 1, 1998, p. 53.

251. N. Burch, *La Lucarne...*, *op. cit.*, p. 224.

sémantique, la présence d'attributs humains dans un dispositif passablement novateur au niveau de sa machinerie. Paul Zumthor souligne que la médiatisation technologique du XX<sup>e</sup> siècle, bien qu'ayant remis la voix sur le devant de la scène, n'en relève pas moins, « en son fondement, [...] de la différence biologique entre l'homme et la machine »<sup>252</sup>. C'est pourquoi on peut s'interroger sur la façon dont cette dichotomie humain/machine a été perçue par ceux qui ont écrit l'histoire du cinéma.

Dans la seconde moitié des années 10, après que le bonimenteur a quitté la plupart des salles européennes et américaines, la voix est relayée en tant que facteur d'humanisation et d'individualisation par le développement de diverses formes de prestations sonores qui furent présentes dès le début de l'exploitation cinématographique, notamment par la musique d'accompagnement. On peut dire qu'à certains égards, la musique a joué le rôle de substitut de la voix des personnages visualisés, convoquant sur la base de conventions culturelles une composante émotionnelle similaire à celle qui transparait, au niveau de la vocalité, dans l'intonation, la diction, le débit ou le grain de la voix (on peut même évoquer la « musicalité » de cette dernière). Cette parenté entre voix et musique est soulignée par Georges Dureau lorsqu'il dit de l'une des premières partitions originales, celle composée par Camille Saint-Saëns pour *L'Assassinat du Duc de Guise* (1908), que cette musique « remplacera la voix humaine dans les moindres détails de son expressivité »<sup>253</sup>. La formulation de ce « critique » contemporain de la création de la Société du Film d'art trahit la nécessité, pour une telle œuvre à prétention littéraire, de trouver un succédané à la voix. Significativement, Dureau ne mentionne pas le verbe, mais le caractère expressif que la source humaine confère à la voix (cette « fonction émotive » induite par l'oralité même). En tant que résultat d'une performance *live*, la musique d'accompagnement est susceptible d'atteindre à l'humanité de la déclamation théâtrale.

Ce lien entre voix et musique s'opère concrètement durant les projections lorsqu'elles sont accompagnées de chants, ou lorsque la musique connaît un « intertexte vocal », ce qui était courant à l'époque. En effet, l'« adaptation » musicale consistait en un mélange

252. P. Zumthor, *Performance...*, *op. cit.*, p. 16.

253. Georges Dureau, « Visions de l'art », *Ciné-Magazine*, 19 novembre 1908, cité par F. Jost, « The Voices of Silence », *art. cit.*, p. 50.

de répertoires préexistants. Or cette pratique, courante dans la France des années 20, visait à intégrer, entre autres morceaux, des «airs à la mode et des succès de music-hall»<sup>254</sup>, c'est-à-dire des mélodies auxquelles les spectateurs associaient inmanquablement une dimension verbale (le texte des chansons) et vocale (la voix d'artistes célèbres). Elle renoue ainsi avec un type d'accompagnement musical plus ancien dont Rick Altman a relevé l'importance dans son étude sur les sons à l'époque des nickelodéons: le pianiste qui accompagnait les projections créait souvent une connivence avec son public en jouant des fragments de chansons populaires dont le titre ou les paroles donnaient lieu, grâce à leur interaction avec le contenu de l'image, à des jeux de mots<sup>255</sup>. Bien que jamais prononcées dans la salle, ces paroles sous-jacentes à la musique étaient donc «actualisées» mentalement par certains spectateurs qui avaient alors le sentiment d'appartenir au même espace culturel que le musicien, ce qui les rapprochait de «l'espace de la production». Lorsqu'il mentionne cet exemple, Altman souligne l'existence de différences individuelles au niveau de l'appréhension de ces sons: «Certains spectateurs connaissaient les paroles de la chanson choisie pour accompagner une scène particulière, et donc comprenaient les liens entre mots et images; d'autres pouvaient apprécier la musique mais rester ignorants de la base linguistique de cet accompagnement.»<sup>256</sup> Ces variations révèlent combien la réception était faiblement standardisée à l'époque du «cinéma» parlé.

La médiation entre l'audiospectateur et la machinerie influe sur la relation horizontale entre l'audiospectateur et la représentation. L'exhibition de la machinerie par l'intermédiaire de la parole tend bien sûr à enrayer les phénomènes de croyance du public envers le monde de l'écran en introduisant un autre objet de persuasion (par exemple relatif à la perfection supposée de l'appareil présenté). On peut donc problématiser l'éventuelle incidence du pôle de la machinerie sur l'illusion référentielle de la représentation. Je propose de traiter ce point à travers quelques discours portant sur le bruit du projecteur, bien que la pertinence de ce facteur soit limitée tant temporellement (peut-être à la période foraine de l'exploitation cinématographique) que géographiquement (on doit distinguer les palaces des centres urbains et les petites salles de province).

254. E. Toulet et Ch. Belaygue, *Musique d'écran...*, op. cit., p. 42.

255. R. Altman, *Silent Film Sound*, op. cit., pp. 220-226.

256. *Ibid.*, p. 281.

Comme il s'agit d'une manifestation sonore, cet indice de mécanique est susceptible d'entrer en concurrence avec la musique et le boniment, et de créer ainsi une dialectique capitale à l'aube d'une ère de «production industrielle des biens culturels». Lorsqu'ils définissent le rôle du *benshi*, Komatsu et Loden soulignent l'importance de l'opposition entre l'humain et la machine :

«Une telle voix humaine était une composante indispensable parmi les conditions <non humaines> de la présentation du film – un produit industriel qui émettait le bruit mécanique d'un projecteur.»<sup>257</sup>

Ainsi, pour ces auteurs, le facteur humain de la voix permet de compenser la présence marquée de la machinerie à laquelle le bruit du projecteur ne cesse de renvoyer. Comme l'a également noté Germain Lacasse, la dimension machinique suscitait une étrangeté que le boniment contribuait à atténuer en créant une certaine familiarité avec le public<sup>258</sup>. Ce pouvoir de la parole n'allait toutefois pas nécessairement de soi: le bonimenteur devait parfois mener une lutte pour maîtriser l'espace sonore, pour se hisser au-dessus des cris du public ou des sons extérieurs à la baraque foraine. Ainsi André Méliès mentionne-t-il le fait que, si le pianiste «devait taper le plus possible pour couvrir le bruit de mitrailleuse de l'appareil», le bonimenteur tentait quant à lui de surenchérir pour «dominer le vacarme de l'appareil et du piano réunis»<sup>259</sup>. Cette évocation, dont la formulation obéit à la hiérarchisation qui prime généralement dans le cinéma parlant (où les bruits sont soumis à la musique, et celle-ci à la voix), rappelle que le bonimenteur était, selon les circonstances, dans la nécessité de s'imposer. On retrouve cette situation ultérieurement lorsque les projections se déroulaient en marge de l'exploitation institutionnalisée. Ainsi en URSS, où d'après Valérie Pozner le nouveau régime semble avoir cherché, dans les années 20, à favoriser le cinéma parlé pour son potentiel didactique, un conférencier-projectionniste donne dans un article de 1926 un aperçu des difficultés auxquelles il est confronté lors d'un boniment simultané aux images en affirmant :

257. H. Komatsu et F. Loden, «Mastering the Mute Image», *art. cit.*, p. 36.

258. G. Lacasse, «Du cinéma oral...», *art. cit.*, p. 53.

259. *Souvenir sur mon père* [1961], cité par Jacques Malthête, «Méliès et le conférencier», *Iris*, N° 22, 1996, p. 118.

«Le bruit de la dynamo est tel qu'il faut forcer sur ses cordes vocales au maximum. Le texte du film doit être hurlé.»<sup>260</sup>

Une autre source de perturbation réside dans les voix qui proviennent du public. Un compte rendu consacré à l'accompagnement d'une projection par le vicomte d'Hauterives témoigne de l'existence d'un tel conflit entre le boniment et les autres voix de la salle :

«À Trois-Rivières, où un groupe de spectateurs chahutait pendant une représentation, imitant le bêlement des moutons vus sur l'écran, le vicomte commente : «Je suis heureux de constater que toutes les bêtes ne sont pas dans le tableau.»<sup>261</sup>

Le bonimenteur se doit d'affermir son autorité sur l'espace occupé par le dispositif et sur le commentaire qui est fait de l'image ; la remarque ironique du vicomte permet de corriger la représentation visuelle et le public, rappelant par l'affirmation du «je» que cette fonction de médiation n'incombe qu'à lui. Généralement, le bonimenteur se distingue passablement du rhapsode ou du prêtre que tous écoutent pieusement : le caractère bruyant du spectacle a forcément des incidences sur son exercice, tandis que le spectateur du cinéma parlant sait que ses propres interventions orales n'affecteraient en rien le déroulement inexorable de la voix fixée. C'est pourquoi la nécessité d'avoir une voix qui porte participe du profil du bonimenteur, dont on exige qu'il soit un véritable «stentor» selon André Méliès.

Comme la voix, la musique semble avoir contribué à occulter le caractère mécanique de la projection cinématographique. Grâce à une intensité sonore bien supérieure à la voix non amplifiée d'un locuteur unique, la musique orchestrale jouée dans la salle – susceptible d'accueillir la voix sous la forme du chant – peut devenir englobante et étouffer les bruits d'origine mécanique. Elle s'est d'ailleurs faite parfois l'ennemi du bonimenteur, étant elle-même le produit d'un degré de «technicité» supérieur, puisqu'elle recourt, littéralement, à des *instruments* autres que le seul appareil phonatoire humain. Ainsi le «conférencier» Georges Dalbe conseillait-il

260. Cité par Valérie Pozner, «Le bonimenteur (rouge). Retour sur la question de l'oralité à propos du cas soviétique», *Cinéma*, Vol. 14, N° 2-3, 2005, p. 164.

261. «Bonnes leçons», *Le Trifluvien*, 11 novembre 1904, cité par A. Gaudreault et G. Lacasse, «L'écran ventriloque», *art. cit.*, p. 141.

en 1911 à ses confrères d'«entretenir les meilleures relations avec les musiciens, car ceux-ci peuvent lui créer bien des ennuis, ils n'ont qu'à jouer fort et le conférencier est tombé»<sup>262</sup>.

Le rôle prépondérant de la musique dans la gestion de l'espace sonore explique pourquoi cette dernière a été l'«instance humaine» la plus discutée dans son rapport aux bruits provoqués par l'appareil de projection. La musique dans la salle de cinéma renoue avec une fonction qu'elle occupait déjà dans le cadre des spectacles scéniques depuis le milieu du XVIII<sup>e</sup> siècle, lorsqu'il lui fallait couvrir, ainsi que l'explique Paul Bergel, les grincements occasionnés par la «machinerie de théâtre» durant les changements de décors :

«On voyait le décor s'ébranler dans le gémissement des poulies et des assemblages. Tout ce vacarme était couvert soit par l'orchestre, soit par des machines destinées à produire des effets sonores tels que le vent ou le tonnerre.»<sup>263</sup>

Au cinéma toutefois, le problème des bruits jugés «parasites» se pose avec plus d'acuité, puisque le bruit du projecteur, intrinsèquement lié à la production même de l'image, est un phénomène constant. Certes, tous les commentateurs n'y ont pas vu un effet néfaste, Céline qualifiant par exemple dans *Mort à crédit* ce son de «ronron du moulin», l'assimilant à une incitation au sommeil ou à la rêverie<sup>264</sup>. Cependant, cette répétitivité mécanique n'est habituellement pas considérée comme un bercement, mais comme une gêne. Bien que cet aspect ne concerne pas en propre la voix, il me semble intéressant d'examiner les arguments avancés par les nombreux auteurs qui ont commenté cette question. Ces discours sont en effet emblématiques de la façon dont on considère communément les liens entre les sons de la salle, l'illusion d'un univers diégétique autonome et la mécanicité du dispositif.

Émile Vuillermoz, auteur de chroniques quotidiennes dans le journal *Le Temps* et critique musical renommé, décrit dans un article de 1917 le rôle de la musique en ces termes :

262. Georges Dalbe, *Le Courrier cinématographique*, 14 octobre 1911, cité dans l'annexe de A. Gaudreault, «Le retour du [bonimenteur] refoulé...», *art. cit.*, p. 32.

263. Cité par Denis Mercier (dir.), *Le Livre des techniques du son*, tome 3 («L'exploitation»), Paris: Éditions Fréquences, 1993 [1<sup>re</sup> édition 1987], p. 221.

264. Louis-Ferdinand Céline, *Mort à crédit*, Paris, Denoël et Steele, 1936, p. 75.



« Beaucoup de spectateurs ne peuvent s'évader dans le rêve, à la suite des fantômes de l'écran, sans être étourdis, bercés et un peu grisés par les vapeurs harmoniques qui montent de l'orchestre et se répandent dans la salle [...]. Le charme serait rompu si le voile des sonorités était brusquement déchiré et si, dans le silence glacial de ce monde des fantasmagories muettes, on ne percevait plus que l'agaçant bourdonnement d'insecte de la machine à explorer le temps et l'espace, qui, murée dans son placard, enroule et déroule sans fin ses télégrammes lumineux. »<sup>265</sup>

Vuillermoz confère à son propos une valeur générale en affirmant que la musique constitue pour certains un tremplin nécessaire à l'immersion dans l'univers filmique. La rupture de l'illusion surviendrait dès lors que le bruit du projecteur se fait entendre, rappelant à l'auditoire que l'image est le produit d'une machine. La vision proposée par le journaliste d'une projection strictement muette s'inscrit dans la droite ligne du texte dans lequel Maxime Gorki rend compte, en 1896, de la façon dont il a perçu les bandes Lumière<sup>266</sup> : il s'agit, pour l'un comme pour l'autre, de « spectres muets ». Dans son ouvrage de 1945 composé de souvenirs, R.-M. Arlaud reconduit cette idéologie naturaliste en expliquant ainsi le rôle du pianiste lors des projections des premiers temps :

« Dès la minute où la photographie s'anima, elle voulut copier la vie et cette imitation *restait bâtarde sans la parole*. Il fallait à tout prix cacher ce mutisme, le maquiller. Dès la première représentation du Salon Indien, il y eut le piano. Pour beaucoup, le piano reste le symbole d'une époque, le symbole d'un âge du cinéma. Ce piano qui, sans répit, sans fatigue, déversait sa musique, ce piano tant moqué, indispensable à la boîte à images, *escamotait autant la sonorité des appareils que le ridicule des muets de l'écran.* »<sup>267</sup>

265. Émile Vuillermoz, « Devant l'écran », *Le Temps*, 18 juillet 1917, cité par E. Toulet et Ch. Belaygue, *Musique d'écran...* *op. cit.*, p. 76.

266. Maxime Gorki oppose le caractère vivant conféré à la représentation par les mouvements des personnes filmées et la mutité qui les condamne à une existence fantomatique : « Les joueurs battent les cartes et... éclatent de rire, de ce rire silencieux des ombres. [...] Et ce rire insonore, ce rire des seuls muscles gris sur ces visages gris agités par l'émotion est complètement fantastique. Un souffle froid s'en dégage qui est par trop éloigné de la vie authentique. Riant, tels des ombres, ils disparaissent, tels des ombres... » (« Le cinématographe Lumière », *art. cit.*, p. 122.)

267. R.-M. Arlaud, *op. cit.*, p. 100 (je souligne).

Selon Arlaud, le piano sert à faire oublier la mutité de la représentation, mais endosse également une seconde fonction – considérée comme réciproque – qui consiste à couvrir le bruit de la machine de projection. On le voit, le silence (des personnages de l'écran) et le bruit (de l'appareil) renvoient à une même déshumanisation d'un spectacle qui exige une intervention machinique. Le pianiste incarne par contre une source sonore à visage humain qui est susceptible d'enrayer chez l'audiospectateur le sentiment déplaisant d'être confronté à la seule mécanicité.

L'évocation des désagréments sonores générés par le mécanisme du projecteur trouvera un écho dans la principale hypothèse avancée par la plupart des historiens pour expliquer l'apparition de la musique d'accompagnement au cinéma dès la projection du Salon Indien. Mentionnée par Arlaud dans le passage cité ci-dessus, elle fut largement diffusée par l'ouvrage *Film Music*, dans lequel Kurt London suggère que la musique aurait été initialement utilisée pour répondre au besoin de « noyer le bruit du projecteur »<sup>268</sup>. On trouve par exemple cet argument chez Albert Laffay et, plus récemment, chez Emmanuelle Toulet<sup>269</sup>. Selon London, l'origine de la musique de film se situe dans les chants des camelots forains, qu'il décrit en faisant référence à *L'Opéra de quat'sous* (*Die Dreigroschenoper*, Georg Wilhelm Pabst, 1930)<sup>270</sup>. En vertu de cet ancrage initial dans la vocalité et la performance, la musique d'accompagnement est donc étroitement liée à l'activité bonimentorielle. Par ailleurs, la discussion du bruit du projecteur participe chez London d'une conception qui présuppose la totale soumission de l'univers sonore au monde de l'image. C'est pourquoi il prône une musique qui ne se fait pas remarquer, afin qu'elle ne perturbe pas le spectateur dans sa concentration sur l'image<sup>271</sup>.

Naturellement, l'explication de London ne tient plus à partir du moment où la sédentarisation de l'exploitation cinématographique conduit, dans les centres urbains, à la construction de « palaces » dotés de cabines insonorisées. Siegfried Kracauer utilise cet argument pour contrer l'idée de London et interpréter différemment la persistance de la musique d'accompagnement<sup>272</sup>.

268. Kurt London, *Film Music*, London: Faber & Faber, 1936, p. 27.

269. Albert Laffay, *Logique du cinéma*, Paris: Masson et C<sup>ie</sup>, 1964, p. 33; Emmanuelle Toulet, *Cinématographe, invention du siècle*, Paris: Gallimard; Réunion des Musées nationaux, 1988, p. 16.

270. K. London, *op. cit.*, p. 25.

271. *Ibid.*, p. 37.

Récemment, Rick Altman a pris le parti de Kracauer, arguant que les autres historiens du cinéma oublient de prendre en compte le développement de projecteurs plus silencieux, ainsi que la généralisation des cabines de projection<sup>273</sup>. Altman fait remarquer que l'argument du bruit du projecteur domine depuis plus d'un demi-siècle les histoires du cinéma et mentionne, outre Kurt London, une myriade d'auteurs anglophones que l'on pourrait inscrire dans le paradigme traité ici. Pour Altman, l'enjeu de cette discussion réside dans la contestation de l'opinion commune voulant que les projections des nickelodéons aient toujours été accompagnées de musique. Il faut néanmoins remarquer que cet état de fait valable pour les États-Unis ne l'est pas nécessairement pour la France. En effet, selon Claude Forest, l'isolation du projecteur en cabine (ou plutôt des deux projecteurs, puisque l'allongement de la durée des programmes avait exigé le dédoublement des appareils) ne s'effectue à grande échelle sur le territoire français que vers la fin de la Première Guerre mondiale<sup>274</sup>. On peut même avancer une date plus tardive, puisque Vuillermoz note dans son article de 1917 qu'en dépit de la séparation instaurée avec l'espace du public (la machine est « murée dans son placard »), le fonctionnement de l'appareil menace de se faire entendre si la musique disparaît. Dans une chronique rédigée en 1923 dans le même journal, Vuillermoz précisera encore qu'« il faut couvrir par un procédé quelconque le petit bruit agaçant de l'appareil de projection »<sup>275</sup>. L'insuffisance de l'isolement acoustique fourni par la cloison est également signalée en 1921 par Ernest Coustet, qui juge nécessaire de « masquer le cliquetis, monotone et énervant à la longue du mécanisme d'entraînement du film, et le sifflement de la lampe à arc ou du chalumeau oxhydrique – bruits qu'atténue un peu, mais ne supprime pas entièrement, la cloison percée qui sépare le projecteur de la salle »<sup>276</sup>.

Au vu d'une situation fort hétérogène qui nécessiterait de nombreux ajustements en fonction de la période historique, du type de salles et de l'emplacement géographique des lieux d'exploitation, je

272. (Note de la p. 138.) Siegfried Kracauer, *Theory of Film, The Redemption of Physical Reality*, London : Oxford University Press, 1960, p. 133.

273. R. Altman, *Silent Film Sound*, op. cit., p. 195.

274. Claude Forest, *Les Dernières Séances. Cent ans d'exploitation des salles de cinéma*, Paris : CNRS, 1995, p. 42.

275. Émile Vuillermoz, « Le synchronisme », *Le Temps*, 9 mars 1923, cité par E. Toulet et Ch. Belaygue, *Musique d'écran*, op. cit., p. 84.

276. Ernest Coustet, *Le Cinéma* [1921], in Marcel L'Herbier, *Intelligence du cinématographe*, Paris : Éditions Buchet-Chastel, 1977 [1<sup>re</sup> édition 1946], pp. 131-132.

ne pense pas qu'il soit pertinent d'épiloguer sur les conditions effectives des projections à cette époque. Dans l'extrait cité ci-dessus, Vuillermoz s'exprime d'ailleurs au conditionnel, manifestant par là une intention plus argumentative que descriptive. Il me paraît par contre intéressant de noter le fait même que l'auteur utilise l'argument du bruit du projecteur, à la fois source d'un visible souhaité et d'un audible gênant, pour suggérer un brusque retour à la réalité. En effet, la mention du bruit du projecteur dans les ouvrages sur le son au cinéma connaîtra une indéniable postérité (notamment chez Michel Chion)<sup>277</sup> qu'illustre de façon particulièrement intéressante l'essai de Theodor Adorno et Hanns Eisler intitulé *Musique de cinéma*. Ces auteurs citent l'hypothèse de Kurt London pour en proposer une interprétation qui, à mon sens, convoque également en filigrane une réflexion sur la voix.

Adorno et Eisler interprètent l'hostilité manifestée par le spectateur envers le bruit de l'appareil de projection comme l'expression d'une crainte qui résulterait de la prise de conscience de la mécanisation provoquée par le dispositif cinématographique. Cette peur « apparentée à la panique », d'ordre collectif, permettrait ainsi de révéler aux masses l'aliénation qu'elles subissent. Les auteurs de *Musique de cinéma* sont d'avis qu'à l'époque du muet, la musique ne se substituait pas à la parole, mais contribuait à atténuer l'impact de son absence, à épargner au spectateur ce qu'ils considéraient, de façon fort similaire à Gorki, comme « le désagrément de voir des copies de personnes vivre, s'agiter et même parler tout en restant muettes »<sup>278</sup>. La perspective des théoriciens interdit tout rapprochement entre la musique et la voix, du moins lorsque celle-ci se manifeste sous la forme individuelle du langage articulé. Ils soulignent certes les avantages de l'ouïe sur la vue en associant la première à un mode perceptif archaïque qui favorise l'intégration de l'individu au collectif; toutefois, la parole me semble participer, dans leur système de pensée, de cet « art bourgeois » dont ils désirent l'éradication par la valorisation d'un *gestus* susceptible de favoriser l'émergence d'un art de masse. Pour Adorno et Eisler, la parole constitue une exacerbation de l'individualisme et de la rationalité propres à l'idéologie dominante. Inversement, la

277. Voir Michel Chion, *La Musique au cinéma*, Paris: Fayard, 1995, p. 38, et *Un art sonore...*, *op. cit.*, pp. 16 et 401-402.

278. Theodor W. Adorno et Hanns Eisler, *Musique de cinéma*, Paris: L'Arche, 1972 [1<sup>re</sup> édition 1947], p. 84.

musique – ou plutôt une certaine musique qu'ils appellent de leurs vœux – permettrait une « prise de distance par rapport à la sphère du privé »<sup>279</sup>. C'est pourquoi ils ont tendance à écarter toute manifestation de la parole, le chant n'étant pas abordé dans sa dimension vocale, mais entièrement rattaché à la musique. Ainsi la parole n'est-elle nullement prise en compte pour l'époque du « muet », ce qui revêt une importance particulière dans la mesure où Adorno et Eisler font de cette absence supposée un argument destiné à valoriser la musique. En effet, l'utilisation de cette dernière dans le cadre de projections cinématographiques se trouve selon eux en quelque sorte naturalisée par un postulat d'antériorité qui en fait une composante ontologique de ce type de spectacle. Ils expliquent que l'on « comprend mieux pourquoi on n'a pas eu l'idée d'utiliser dans les débuts du cinéma le procédé apparemment le plus facile : accompagner le film de dialogues confiés à des récitants dissimulés comme dans le théâtre de marionnettes ; on n'a pensé qu'à la musique qui pourtant n'avait aucun rapport avec l'action dans les vieux films d'épouvante et les farces »<sup>280</sup>. L'état actuel des recherches en histoire du cinéma permet d'affirmer le contraire, puisqu'il est désormais avéré que la pratique des locuteurs cachés derrière l'écran a eu cours et qu'elle s'est professionnalisée dans les années 10, par exemple sous l'impulsion de Georges Lordier en France ou, surtout, aux États-Unis grâce aux compagnies de vaudeville. D'ailleurs, l'une des premières projections des frères Lumière recourait déjà à ce principe, la bande intitulée *Discussion de M. Janssen et de M. Lagrange* ayant été projetée le 11 juin 1895 alors que les deux intéressés récitaient le dialogue derrière l'écran<sup>281</sup>. Favorisée dans les premières années du cinématographe par le système de la projection par transparence<sup>282</sup>, cette pratique du doublage *live* s'est ensuite poursuivie, apportant un gain de spectacularité aux séances. Il est vrai néanmoins qu'elle ne s'est pas institutionnalisée dans des proportions comparables à la musique d'accompagnement.

Dans une logique similaire, Adorno et Eisler taisent la présence du bonimenteur, bien que celui-ci soit mentionné par Kurt London

279. *Ibid.*, p. 20.

280. *Ibid.*, p. 85.

281. Voir Bernard Chardère et Guy et Marjorie Borgé, *Les Lumière*, Lausanne : Payot, 1985, p. 73.

282. Voir Cl. Forest, *op. cit.*, p. 29.

qui, dans le paragraphe précédant celui qu'ils citent, met en relation cet énonciateur avec le bruit de la mécanique de projection («Le projecteur faisait un bruit terrible: le bonimenteur [commentator] devait avoir de bons poumons pour rendre son discours audible») <sup>283</sup>. On ne peut certes leur faire grief d'une méconnaissance d'aspects historiques qui n'étaient pas du tout discutés à leur époque; il est cependant évident que l'asservissement de telles données historiques à leur discours théorique participe d'une démarche visant à rejeter les dimensions verbale et vocale. Cette perspective les conduit également, dans le bref passage qu'ils consacrent à la question de la parole et du dialogue au cinéma, à souligner l'inadéquation du couplage de la voix et de l'image en arguant que l'«irréalité» de celle-ci s'accommode mal de la fidélité de la reproduction de celle-là. Leur position repose en partie sur l'argument traditionnel de la spécificité visuelle du médium – ils ne manquent d'ailleurs pas de préciser que «tout ce qui parle au cinéma a quelque chose *d'impropre*» <sup>284</sup> – mais s'avère singulière en ce qu'elle postule une certaine *déliaison* entre la voix et l'image. Les auteurs de *Musique de cinéma* sont sensibles au caractère paradoxal de la représentation audiovisuelle, qui imite la réalité alors que «les mots sortent d'une bouche immatérielle». Afin de valoriser l'apport de la musique, Adorno et Eisler font l'impasse sur les différences existant entre le cinéma muet et le cinéma parlant. Selon eux, «l'insistance dont témoigne le cinéma parlant à se donner pour le double sans faille du monde extérieur dans sa totalité est perçue comme imposée par la ruse et comme fragile» <sup>285</sup>. Une telle remarque, révélatrice de l'obstination des auteurs à considérer la synchronisation de l'image d'un locuteur avec sa voix comme un phénomène bancal, peut sembler quelque peu anachronique dans un ouvrage écrit plus de quinze ans après la généralisation du cinéma parlant aux États-Unis. En niant l'importance d'un mode de représentation qui s'est imposé de façon hégémonique, Adorno et Eisler se démarquent radicalement de l'avis dominant à leur époque pour renouer avec des considérations émises à la fin des années 20, par exemple chez Luigi Pirandello, pour qui «la voix vive instaure un contraste insurpassable avec [la] qualité d'ombre [des images]», de sorte qu'elle «dénonce

283. K. London, *op. cit.*, p. 27.

284. Th. W. Adorno et H. Eisler, *op. cit.*, p. 85.

285. *Ibid.*, pp. 85-86.

le mécanisme»<sup>286</sup>. Il est cependant étonnant que ces partisans d'une représentation «danciée» au sens de Brecht – dont ils se réclament explicitement<sup>287</sup> – ne valorisent pas la «disparité technique entre l'image et le mot»<sup>288</sup>. Au contraire, ils semblent paradoxalement dénigrer la voix synchrone en raison de son inaptitude à créer le «double sans faille du monde extérieur», en accord avec ce qu'ils ont défini au début de leur opus où ils affirment que l'objectif du cinéma est de «représenter avec exactitude la réalité»<sup>289</sup>. Contrairement aux autres auteurs mentionnés jusqu'ici, Adorno et Eisler ne voient pas dans la pratique dominante de la voix synchrone un facteur d'occultation de la mécanicité du dispositif audiovisuel. À leur avis, la musique contribue par contre à atténuer le choc produit par la rencontre de la mécanique et de l'humain. Par conséquent, ils font de l'aspect mécanique de la projection un interprétant essentiel de l'analyse des sons au cinéma.

L'explication selon laquelle la musique d'écran aurait répondu au besoin de couvrir le bruit de l'appareil s'accompagne systématiquement d'une prise en compte de la réception spectatorielle. Edgar Morin, pour qui la musique constitue avant tout un facteur de participation affective, accepte cet argument tout en supposant qu'il participe de quelque chose de plus fondamental :

«Et pourtant la musique s'est imposée au film [...]. Est-ce pour dissimuler les ronflements de l'appareil de projection comme le veut une légende tenace? (Un incident contingent explique toujours ce dont la logique nous échappe.) Pourquoi pas d'ailleurs? L'important est que la musique ait débordé, et de loin, ce rôle protecteur.»<sup>290</sup>

En 1958, la thèse du lien entre l'émergence de la musique et le bruit de l'appareil fait donc figure de «légende tenace». Morin ne la contredit pas, mais n'en tire pas de conclusions quant au rôle joué par la musique dans la constitution de l'imaginaire offert au spectateur.

286. Luigi Pirandello, «Se il film parlante abolirà il teatro» [1929], cité par Alberto Boschi, «Il cinema sonoro nella teoria classica», in *L'Immagine acustica II, Cinegrafie*, N° 6, 1993, p. 66.

287. Th. W. Adorno et H. Eisler, *op. cit.*, p. 34.

288. *Ibid.*, p. 75.

289. *Ibid.*, p. 15.

290. Edgar Morin, *Le Cinéma ou l'homme imaginaire*, Genève: Gonthier, 1965, p. 69.

Rick Altman, dans son ouvrage sur la comédie musicale – un travail antérieur à ses recherches plus approfondies sur les « sons du muet » qui le conduiront à nuancer ou à rejeter certains lieux communs énoncés à propos de ces questions – reconduit l’argument de l’influence du bruit du projecteur sur la représentation en se référant de façon indifférenciée à l’ensemble de la période muette :

« Venant de derrière, un ronronnement, un grincement mécanique rappelle à tout instant que le film n’est pas la réalité, qu’il n’est que le produit d’une projection. Non loin de l’écran, un pianiste actionne les touches en s’évertuant désespérément à couvrir le bruit du projecteur, à faire oublier l’artificialité de l’image : le projecteur est au piano ce que la réalité est au cinéma. »<sup>291</sup>

Altman met l’accent sur une dichotomie située au niveau de « l’espace », dans son sens littéral (celui de la salle, partagé entre l’avant et l’arrière) et métaphorique (relatif à l’immersion du spectateur dans l’espace fictif du film). La musique se trouve du côté de l’imaginaire (et concrètement à proximité de l’image projetée), alors que le projecteur participe de la machinerie des coulisses qu’il s’agit d’occulter. En spéculant sur les intentions d’un pianiste de l’époque et leur prétendue inanité, Altman veut souligner combien il était nécessaire de faire taire le bruit mécanique. Cet exemple peut s’appliquer également au bonimenteur puisqu’il était aussi situé, selon les cas, dans l’espace jouxtant l’écran. Altman effectue ensuite une « transposition » dans le film parlant où une même opposition fonderait la distinction entre les sons diégétiques et les sons extradiégétiques. Le passage au parlant s’accompagne dans le discours de l’historien d’une prise en compte implicite de la voix, puisqu’il est désormais question de la « piste sonore » dans son ensemble. Cette transposition s’avère problématique si l’on compare les deux couples de termes : la musique extradiégétique ne peut être au film parlant ce que le bruit de projecteur fut au muet, puisqu’il incombait au pianiste précité de fournir une musique d’accompagnement qui, dans un très grand nombre de cas, ne pouvait concorder avec une source visualisée à l’écran. Cette mélodie au piano est donc plutôt assimilable, du moins dans ses rapports à la

291. Rick Altman, *La Comédie musicale hollywoodienne*, Paris : Armand Colin, 1992 [1<sup>re</sup> édition 1987], pp. 76-77.



diégèse filmique, au statut d'«extradiégéticité» de la musique du parlant. Bien que maladroit, le rapprochement d'Altman montre combien le bruit du projecteur constitue une référence idéale lorsqu'il s'agit d'opposer l'artifice au «réalisme». Le facteur humain joue selon moi un rôle considérable; Altman personnalise d'ailleurs la musique en évoquant le pianiste. Ainsi une voix provenant de l'espace de la salle, en dépit d'importantes déficiences en termes de synchronisme – ou peut-être en vertu même de ces carences que l'on oppose à la perfection de la machine – contribue-t-elle également à atténuer la dimension mécanique de la projection, et donc à favoriser la «suspension de l'incrédulité» du public face aux personnages de la fiction du film. La postérité de cette conception du rôle des sons durant les projections «muettes» s'explique par le sentiment, ressenti ou imaginé par les théoriciens, d'une inadéquation du tout-mécanique avec le monde humanisé de l'écran. En tant que médiateur, le bonimenteur facilite la corrélation de ces deux pôles du dispositif.

On peut se demander ce qu'il advient de cet objet de réflexion durant la période cruciale de la généralisation des sons fixés. Un article rédigé en 1928 par Jean Prévost à propos du film *Ombres blanches* (*White Shadows in the South Seas*, W. S. Van Dyke et Robert Flaherty, 1928) apporte un éclairage intéressant. L'auteur y explique comment il perçoit les nouveaux liens qui s'instaurent entre la musique et le bruit de l'appareil :

«Supposons que cette musique [enregistrée et synchrone] sera toujours fournie par un meilleur orchestre que celui qui jouerait dans la salle: restera le manque de velouté; restera aussi que cette musique ne peut donner aucun moment de silence. L'orchestre ordinaire a pour fonction essentielle de voiler le bruit mécanique du film qui se déroule; le public est censé ne pas aimer ce bruit, qui me plaît et me rappelle le claquement du vent de mer dans les voiles. Mais l'orchestre mécanique met la machine en mouvement, et commence donc par produire ce bruit qu'il a pour fonction de voiler.»<sup>292</sup>

Notons que le film dont il est question, *Ombres blanches*, ne comporte que quelques rares paroles synchrones – les personnages qui

292. Jean Prévost, «L'étudiant de Prague; Ombres blanches», *La Nouvelle Revue française*, tome XXXII, 1929, p. 130.

s'expriment oralement étant fréquemment vus de dos (à moins qu'ils ne communiquent, comme les amants, en imitant des gazouillements d'oiseaux), alors que les chants sont rapportés à des instances collectives dans lesquelles il est impossible d'individualiser une source visible – la majorité des dialogues étant transmise par l'intermédiaire d'inter-titres; par contre, la musique et les effets sonores sont omniprésents. Prévost aborde donc ce film «sonore» – qui s'inscrit selon Alexandre Arnoux (dans un article de *Pour Vous* du 31 janvier 1929) parfaitement dans la filiation des accompagnements du muet<sup>293</sup> – en opposant orchestres «ordinaire» et «mécanique»: le second surpasse le premier en termes de qualité de l'exécution musicale (alibi culturel qui fut un argument promotionnel central des compagnies de production), mais implique une perte de «velouté». Ce terme fait référence au «sensorium» éveillé par une performance vivante (on parle aussi du velouté de la voix et du regard) – en l'occurrence celle des musiciens présents dans la salle de cinéma, c'est-à-dire à une caractéristique foncièrement humaine. Pour Prévost, la médiation dans la salle constituait un facteur de rapprochement entre l'orchestre et les spectateurs, la garante d'une certaine «tactilité» que la sonorisation «fixée» évince partiellement. Le critique mentionne ensuite la fonction de recouvrement du bruit du projecteur, endossée jusque-là par la musique d'accompagnement. Son article témoigne de la persistance de l'argument du bruit du projecteur dans la réflexion ultérieure à l'isolation en cabine des appareils. En jouant sur la proximité étymologique entre le verbe «voiler» et le son agréable des «voiles», Prévost suggère que les deux termes du couple oppositionnel occultation/exhibition peuvent être valorisés. La séance d'*Ombres blanches* a révélé au critique que la musique phonographiée ne provoque pas l'effet de dissimulation du pôle machine qu'atteignait la musique d'écran, mais conduit au contraire à une exacerbation de la dimension mécanique. C'est que l'illusion exige de la part de l'auditeur une occultation des bruits de fond indissociables du fonctionnement du gramophone. Le médium technique produit inévitablement son propre souffle, indice d'une matérialité technique dont est empreinte la vocalité de la parole fixée.

Notons que l'appréhension de cette situation par les contemporains de la sortie du film sonorisé de Van Dyke n'est pas généralisée. Le critique Georges Auric affirme quant à lui en décembre 1928 que

293. Cité par Roger Icart, *La Révolution du parlant, vue par la presse française*, Perpignan: Institut Jean Vigo, 1988, p. 221.

la «suppression de la fosse d'orchestre [...] permet une projection tout à fait directe et proche du spectateur»<sup>294</sup>. Avec Prévost et Auric, nous sommes en présence de deux conceptions antagonistes du son au cinéma et du rapport entre la médiation vivante et la machine: Auric, à l'instar de nombreux critiques spécialisés du moment, prône le sonore contre le parlant, alors que le rédacteur de la *NRF* promeut exactement l'inverse («le film intégralement parlant ou rien»). Ces divergences sont révélatrices des deux paradigmes (l'un anthropocentrique, l'autre techniciste) qui ont traversé la réception de l'usage du son au cinéma. On trouve d'une part l'oralité liée à des pratiques qui eurent cours durant le «cinéma» parlé, d'autre part la fixation, modelée sur l'écriture, qui trouve son point d'aboutissement dans les différentes technologies du cinéma sonore ou parlant. Lors de la projection d'*Ombres blanches*, film produit par la MGM et réalisé avec le système de «son sur film» Movietone, l'émission même de la musique est directement liée à l'entraînement de la pellicule. Si, pour Prévost, la musique mécanique est incapable de rendre le silence, c'est qu'un bruit constant émane de l'appareillage même utile à la reproduction des sons. Ainsi l'auteur relève-t-il, en rapprochant le sujet du film des conditions technologiques de sa réalisation, le paradoxe souligné par *Ombres blanches* dans lequel on fait «traduire à cette invention nouvelle justement la haine de la civilisation et du mécanisme»<sup>295</sup>. Probablement le statut passablement hybride du film de Van Dyke – tant au niveau de l'image qui fut le résultat d'une collaboration avec Flaherty qu'à celui de la sonorisation partielle décidée une fois le tournage effectué – n'est-il pas pour rien dans les réactions de Prévost; toutefois, ses remarques atteignent un degré de généralité qui touche au principe de la sonorisation proprement dit. Ce qu'il énonce à propos de la musique ne semble cependant pas valable pour la voix, puisqu'il prône le 100% parlant contre l'esthétique d'*Ombres blanches*. Il ne donne aucune précision sur ce point, mais on peut imaginer que le statut exceptionnel de la voix a trait à l'humanité fondamentale qui lui est associée.

La mécanicité intrinsèque de la reproduction sonore soulignée par Prévost représente un paramètre central de la réception des premiers films parlants et des enregistrements phonographiques, même si l'évolution technologique et les habitudes perceptives du public

294. Cité par *ibid.*, p. 178.

295. J. Prévost, *art. cit.*, p. 131.

tendent à rendre cette mécanique imperceptible. En plus des imperfections de la technique de diffusion des sons dans la salle, le cinéma parlant introduit une nouvelle donnée comparable au cliquettement du projecteur, mais liée à la genèse du film : il s'agit de bruits parasites occasionnés par la prise de son qui, jusqu'à la fin de l'année 1929 environ, ne permettait le synchronisme avec l'image qu'à la condition d'avoir lieu simultanément au filmage. Rick Altman note à propos de l'introduction du *play-back* dans la réalisation de comédies musicales « qu'auparavant l'orchestre et les chanteurs étaient enregistrés sur fond de ronronnement de caméra, de bruits de pas de figurants, de toussotements de l'assistance »<sup>296</sup>. Cette formulation se modèle totalement sur le discours dominant relatif au bruit du projecteur en postulant qu'avec l'enregistrement sonore, c'est aussi l'appareil de prise de vues qui se manifeste et rappelle constamment l'origine technique des images. La postsynchronisation joue alors un rôle semblable à la musique d'accompagnement des bandes muettes. À l'instar des auteurs qui ont abordé les bruits dans la salle, Altman semble sous-estimer les moyens mis en œuvre pour atténuer les nuisances sonores provoquées par le fonctionnement de l'appareil de prise de vues. Sur la base de photos de plateau, Martin Barnier a montré que les caméras ont rapidement été affublées, sur le tournage de *talkies*, d'un caisson insonorisant<sup>297</sup>. Toutefois, ce qui est valable pour les productions hollywoodiennes du début des années 30 ne l'est pas nécessairement pour toutes les cinématographies ; par exemple, dans *Le Chant de la fleur écarlate*, film finlandais réalisé en 1938 par Teuvo Tulio et obéissant à une esthétique encore fort redevable, dans certaines séquences, de l'époque muette ou de la phase de transition des films sonores dépourvus de paroles, on entend tout à fait distinctement le moteur de la caméra dans les plans rapprochés comprenant des dialogues.

Il ne s'agissait pas, dans ce chapitre, de contribuer à l'écriture d'une histoire des techniques, mais à s'interroger sur des discours dont l'objet fait intervenir les pôles machinique et spectatorial. Or, on constate que la question de l'isolation phonique, propice à une réflexion sur la mécanique du dispositif, fut une préoccupation tant pour les réalisateurs de films, pour les exploitants que pour les théoriciens du cinéma.

296. R. Altman, *La Comédie musicale hollywoodienne*, op. cit., p. 168.

297. M. Barnier, *En route...*, op. cit., pp. 76-77.

## II.5 UNE SEANCE MELIÈS BONIMENTEE PAR MADELEINE MALTHÊTE-MELIÈS

Caractérisée par l'absence de fixation et l'importance du *hic et nunc* de son déroulement, la voix vive du boniment pratiqué durant la période du cinéma des premiers temps n'a pu nous parvenir que sous la forme de témoignages, ou par l'intermédiaire de documents écrits qui sous-tendaient parfois le texte proféré. Or, si l'on s'appuie sur des sources écrites pour envisager une situation d'oralité (et à plus forte raison pour tenter de la restituer), on perd ce qui fait les fondements de cette dernière. D'ailleurs, même si nous étions en présence d'un bonimenteur du début du siècle, la réception que nous aurions du spectacle ne permettrait en aucun cas de rendre compte de ce qu'elle fut à l'époque. André Gaudreault explique ainsi les raisons de cet insurmontable obstacle: «Il est impossible de restaurer un «contact» film-spectateur qui date d'une époque antérieure qui, donc, se situe dans un autre paradigme culturel.»<sup>298</sup> Cette différence de posture perceptive vaut également pour les témoignages postérieurs de spectateurs contemporains du «cinéma» parlé. En effet, lorsqu'ils rédigent leurs souvenirs, ils n'échappent pas non plus à l'influence exercée par les nouveaux paradigmes qu'ils ont inconsciemment intégrés.

Ces réserves n'ont pas empêché André Gaudreault de participer, entre 1990 et 1997, à trois «restitutions» de ce type. Une reconstitution peut en effet offrir un «laboratoire» utile pour une réflexion menée sur la médiation bonimentorielle, notamment sur le plan théorique du dispositif du «cinéma» parlé, pour autant que l'on fasse preuve d'une grande prudence lorsqu'il s'agit de formuler des hypothèses d'ordre historique. C'est dans cette optique (et avec de telles précautions) que j'ai décidé de clore ce chapitre consacré au boniment par l'analyse d'une séance de projection qui s'est déroulée au Théâtre du Musée Grévin en 1997. Il s'agit d'un programme de films de Georges Méliès bonimenté par la petite-fille du cinéaste, Madeleine Malthête-Méliès. Bien que cette pratique se présente comme un «héritage familial», il ne faut pas se méprendre quant à l'exactitude de la reconstitution, manifestement réalisée pour un public d'aujourd'hui et dans l'intention de perpétuer la mémoire du pionnier. En outre, cette séance ne peut être étudiée dans sa dimension événementielle, mais doit être

298. A. Gaudreault, «Fonctions et origines...», *art. cit.*, p. 120.

considérée en tant que trace, telle qu'elle existe sous la forme d'une édition DVD (*Méliès le Cinémagicien*, Arte Vidéo, 2001). L'objet n'est donc pas réduit au seul boniment, devenu événement profilmique, mais au film même qui en rend compte et offre à ce dernier une diffusion beaucoup plus large. L'examen du film *Une séance Méliès* (Jacques Mény, 2001), qui comprend quinze métrages du « cinéaste » français et dure cinquante-cinq minutes, engage trois niveaux :

- celui des bandes projetées proprement dites, accompagnées d'une voix qui est alors systématiquement *off*;
- celui du contexte profilmique plus large dans lequel s'inscrivent les conditions de la projection au Musée Grévin;
- celui de la restitution filmique elle-même, qui induit des transformations du continuum profilmique de la séance, ne serait-ce que parce que celle-ci, apparemment ininterrompue à l'écran, s'est en fait déroulée sur deux jours, le cinéaste précisant dans un article que le tournage eut lieu les 29 et 30 avril 1997<sup>299</sup>.

En faisant fusionner deux temps dans une unité fictive, le film extrait le représenté de son ancrage dans le présent qui, comme on l'a vu, constitue un aspect déterminant de la performance. L'organisation liée à la restitution filmique intervient au niveau du montage, qui fait majoritairement alterner, après un prologue au cours duquel le public prend place pendant que la séance est annoncée par un bonisseur muni d'un orgue de Barbarie, trois types de plans qui correspondent à autant d'instances de l'espace de production : outre les films de Méliès, on nous montre la bonimenteuse qui, lors d'intermèdes où la lumière est rétablie dans la salle, s'adresse frontalement à la caméra; le pianiste, le plus souvent réduit par synecdoque à ses mains filmées en gros plan sur le clavier à l'issue de la projection d'une bande; enfin, l'appareil de projection qui, actionné manuellement, est exhibé juste avant le début de certains films du programme. Notons que la présence de cette instance – dont la mécanique est soulignée par le fait que le projectionniste demeure dans l'ombre – se manifeste également au niveau sonore, puisque le bruit du projecteur est audible. On a vu que cette composante sonore avait donné lieu, à l'époque du « muet », à des réflexions qui

299. Jacques Mény, « Méliès imaginé ou images de Méliès au cinéma et à la télévision », in *Georges Méliès, l'illusionniste fin de siècle?*, *op. cit.*, p. 382.

engageaient la place de l'humain dans le dispositif. Or la représentation filmique qui nous en est donnée dans *Une séance Méliès* s'inscrit tout à fait dans les exigences qui ont été formulées à son égard : retentissant distinctement lorsque l'appareil est filmé en plan rapproché, ce bruit « parasite » est ensuite étouffé grâce à de rapides fondus sonores dès que les films de Méliès apparaissent à l'écran. Il est à la fois un rappel des conditions de projection de l'époque et un élément dont l'occultation est nécessaire afin que les usagers du DVD puissent jouir pleinement des œuvres de Méliès. Ce travail sur le son met en évidence le caractère construit du film de Mény : il s'agit de la trace d'une séance qui, elle-même, se veut la trace d'une pratique passée.

Bien que les points évoqués restreignent la portée de l'analyse du boniment, on peut dégager certaines tendances générales en étudiant les rapports entre la parole et la représentation visuelle. Tout d'abord, Madeleine Malthête-Méliès évoque explicitement les trois instances responsables de la gestion du spectacle (le projectionniste, le pianiste et elle-même), de sorte qu'elle se positionne en tant que médiation entre le spectateur et la machinerie. Cette dominante fonctionnelle de type autoréférentiel se manifeste par la suite régulièrement lorsque la bonimenteuse parle de Méliès en tant que cinéaste et considère les films sous l'angle de leur réalisation. Il s'agit pour elle d'adapter le dispositif à la culture-cible. Or cette démarche implique de combler les lacunes du spectateur de la fin du XX<sup>e</sup> siècle relatives aux conditions de projection de l'époque ou à certains référents de l'image qui ne peuvent plus être établis aujourd'hui, à l'instar des marques de produits disposés dans le champ à des fins publicitaires (le champagne Le Mercier dans *Barbe-Bleue*) ou parodiques (*Les Affiches en gogouette*).

On constate cependant que la fonction principale de la bonimenteuse consiste à clarifier verbalement le récit visuel. La brièveté des bandes est cependant souvent l'indice d'une absence de développement narratif. C'est pourquoi nous ne trouvons aucun boniment simultané aux images dans la quasi-totalité des films uniponctuels (parmi les quinze films présentés, seuls trois comprennent plus d'un plan)<sup>300</sup>, qui se caractérisent par un contenu

300. J'entends le terme « plan » au sens d'un changement perceptible de décor ou d'une ellipse temporelle lorsque ces modifications sont imputables au « montage ». Il est certes évident qu'il y a collure lorsque certains trucages sont effectués à l'intérieur même de ce que j'appelle un « plan » ; inversement, la transformation des décors grâce à une machinerie scénique n'engage pas le montage. On sait par ailleurs que Méliès envisageait le « découpage » de ses films en fonction d'une unité qu'il qualifiait de « tableau ».

narratif fort ténu ou nul. En effet, lorsque ces films à un seul plan font reposer l'intérêt sur la valeur attractionnelle propre aux actions montrées et aux trucages, qu'il s'agisse de combats (*Nouvelles luttes extravagantes*), de tours de prestidigitation (*Un Homme de tête, Le Mélomane, Le Thaumaturge chinois, Les Cartes vivantes*) ou de danse (*Le Cake Walk infernal, Le Chaudron infernal*), la bonimenteuse «laisse parler» les images. Ce principe vaut pour l'ensemble de la séance, puisque aucun trucage (comme les surimpressions, les effets pyrotechniques ou les arrêts de caméra) n'est souligné verbalement, que ce soit au niveau autoréférentiel (où le procédé pourrait être commenté en tant que tel) ou référentiel (les actions résultant des trucages ne sont jamais mentionnées). Cette absence de mise en évidence des effets visuels est d'autant plus notable que ceux-ci favoriseraient l'établissement de points de synchronisation, puisque les actions représentées (apparitions, disparitions ou explosions) jalonnent le déroulement de la bande-image sur un mode ponctuel. La bonimenteuse opte néanmoins pour une approche foncièrement linéarisante, alors que l'exhibition des «clous» du spectacle s'inscrirait au contraire dans une logique de la fragmentation. Le silence de la bonimenteuse durant les trucages assure par contre une réception *immédiate*, tandis que le texte risquerait d'enrayer l'effet de surprise ou le pouvoir d'illusion des effets obtenus sur le plan visuel.

Ce choix nous permet de comprendre pourquoi seulement deux des films uniponctuels sont pourvus d'un commentaire simultané visant à remédier à d'éventuelles lacunes du public actuel en recourant aux fonctions référentielle et réflexive. Il s'agit d'une part des *Affiches en goguette*, bande évoquée ci-dessus, d'autre part du *Tripot clandestin*, où la bonimenteuse précise une action puis ses conséquences. Dans ce second film, le texte pointe de façon synchrone les deux phases-clés du récit, l'extinction de la lumière et le moment où quelqu'un la rallume. Cette narrativisation n'est pas dénuée d'une certaine autoréférentialité: c'est aussi le «code» de l'opposition jour/nuit issu des spectacles scéniques que le boniment précise pour le spectateur d'aujourd'hui qui, même s'il a coutume de la «nuit américaine», pourrait ne pas comprendre que le combat se déroule dans l'obscurité.

Si elle se fait un peu plus loquace dans ces deux bandes, Madeleine Malthête-Méliès se contente, dans la grande majorité des projections, de proposer un bref commentaire qui, antérieur à la



projection, se termine par la mention du titre du film. Cette importance accordée aux titres me paraît essentielle pour saisir les options « traductionnelles » qui caractérisent ce boniment, et qui sont sans doute le fruit d'une volonté de respecter les indications conservées dans les catalogues où les différents « tableaux » étaient répertoriés par titre. La notion de « titre » est en effet inséparable de certaines fonctions du verbal. Pour Leo H. Hoeck, il s'agit d'un ensemble de « signes linguistiques [...] qui peuvent figurer en tête d'un texte pour le désigner, pour en indiquer le contenu global et pour allécher le public visé »<sup>301</sup>. Quel que soit le moyen d'expression concerné, trois rôles peuvent lui incomber : une fonction désignative (c'est là l'unique trait définitoire qui s'avère obligatoire), une fonction narrative dont le principe est celui de la condensation et, en tant que première information livrée au lecteur ou au spectateur, une fonction pragmatique. La fonction désignative est d'autant plus importante à l'époque des projections bonimentées qu'en l'absence de générique, le titre constituait un élément de reconnaissance majeur pour l'exploitant et le public, même s'il n'était pas associé, comme il l'est de nos jours, aux droits sur la propriété intellectuelle, ainsi qu'en témoignent les difficultés rencontrées (notamment par Méliès) pour protéger les films des contre-types vendus par la concurrence. Lorsque Gaudreault discute les postulats « proto-narratologiques » qui informèrent l'élaboration d'un copyright pour les films, il fait remarquer que les cartons introduits dans les films d'Edison faisaient avant tout office de titres, leur fonction désignative permettant que les différents tableaux fussent enregistrés de façon indépendante sur le plan légal<sup>302</sup>. On a d'ailleurs dénommé « titres » ces mentions écrites insérées dans les films durant toute l'époque muette, ce terme valant comme synonyme d'« intertitre » ou de « sous-titre » avant qu'une terminologie plus précise ne se généralise avec le parlant.

L'attractivité du titre – souvent liée au caractère attractionnel des sujets annoncés – concernait en priorité le bonisseur qui haranguait la foule à l'extérieur des lieux de projection, mais aussi le bonimenteur qui se devait de captiver le public afin qu'il ne quitte pas la salle. Dans les dénominations des films de Méliès, ce sont surtout les adjectifs qui jouent ce rôle d'accroche (« extravagant »,

301. Leo H. Hoeck, *La Marque du titre. Dispositifs sémiotiques d'une pratique textuelle*, La Haye; Paris: Mouton, 1981, p. 17.

302. A. Gaudreault, *Du littéraire...*, *op. cit.*, p. 128, note 16.

«infernale», «diabolique», «vivant», «clandestin», et peut-être même «chinoise»). Par contre, les titres présentent pour la plupart un faible potentiel d'extension narrative, à l'exception notable de *Voyage dans la lune*, et, grâce à l'intertexte qu'il convoque, de *Barbe-Bleue*. Or ces deux films sont justement les plus longs du programme, et ceux dont la projection filmée par Mény connaît le plus d'interventions de la part de la bonimenteuse.

De façon générale, Madeleine Malthête-Méliès se fait donc plutôt discrète : ses énoncés, simultanés aux projections, sont majoritairement fort brefs. De plus, elle ne fait montre d'aucune performativité particulière dans l'usage de la voix. Ce retrait, souligné par le fait qu'elle s'assoit dans l'obscurité parmi le public lors des projections, s'explique probablement par une volonté de laisser la part belle à son aïeul auquel la séance rend hommage, comme le rappelle un portrait fixe du cinéaste projeté entre chaque film. Cette posture bonimentorielle n'interdit toutefois pas que Madeleine Malthête-Méliès joue un rôle important dans la gestion de la séance. En effet, sa voix constitue un liant qui traverse des films parfois fort brefs (les deux tiers d'entre eux durent moins de trois minutes) et fait du programme dans son ensemble – bien plus que des films en particulier – une unité. L'entracte parlé durant lequel les rideaux rouges se ferment devant l'écran alors que la bonimenteuse prend place sur la rampe pour raconter une anecdote, les présentations avant chaque projection et l'encadrement du programme déterminent la structure de la séance. L'autonomisation du «montage» de la séance par rapport à l'individualité des films est renforcée par le fait que ceux-ci ne sont pas présentés dans un ordre induit par leurs caractéristiques propres : on ne respecte ni la chronologie, ni un découpage thématique, fût-il établi sur de vagues parentés. Ce qui prime, c'est l'organisation globale de la séance, son rythme général. Il est en effet communément admis aujourd'hui chez les historiens du cinéma que l'unité qui prévaut à l'époque du cinéma des premiers temps est le programme ou la séance<sup>303</sup>. Placé sous le régime de l'oralité, le «film» n'acquiert pas l'autonomie discursive qui s'affirmera au cours de la première moitié des années 10, alors que les bonimenteurs commençaient à se faire rares. Il est à cet égard significatif que les deux métrages les plus longs du programme bonimenté par Madeleine Malthête-Méliès,

303. Voir François Albera, «Du bonimenteur à la voix off», *Vertigo*, N° 26, 2004, p. 46.

*Barbe-Bleue* (1901) et *Le Voyage dans la lune* (1902), ne soient pas projetés consécutivement (ils occupent respectivement l'antépénultième et la seconde place), mais fassent figure de « clous » au sein des parties initiale et finale du film de Mény. Afin d'examiner plus précisément les liens qui s'instaurent entre le boniment et les images, je vais me concentrer sur ces deux films dans lesquels la présence de la médiation vocale est forte.

Le premier plan du *Voyage dans la lune* se caractérise par une forte densité de protagonistes et d'actions. Dans ces circonstances, le boniment s'évertue surtout à clarifier cette exposition en sélectionnant les principaux informants narratifs et en individualisant les personnages. Cette seconde opération passe par la nomination du héros (« le professeur Barbenfouillis ») – présentation synchrone avec la révérence de ce dernier à l'écran – et une précision concernant ses collègues : « enfin, cinq courageux héros vont accepter de le suivre dans sa folle entreprise ». La mise en évidence des futurs voyageurs renforce l'individualisation qui s'effectue également à l'image lorsque les personnages s'avancent au premier plan et changent de costume. Notons que ces cinq personnages sont même nommés dans le catalogue d'époque : il incombait au verbal de fournir des repères sous la forme d'éléments constants<sup>304</sup>. La mention du seul Barbenfouillis par la bonimenteuse introduit ici une hiérarchisation, probablement liée au fait que ce rôle est tenu par Méliès lui-même. Le futur périphrastique (« vont accepter ») permet de concilier le rapport de proximité de la médiation et la fonction d'annonce, qui prévaut dans la plupart des cas (ne serait-ce que dans tous les « titres »). Le lexique (courageux/fou) à la fois emphatique, naïf et humoristique (« Barbenfouillis »), évoque l'atmosphère des récits populaires et l'exigence d'une adhésion inconditionnelle de l'auditeur, happé dans la diégèse dès la première phrase par le pronom « nous ». Ne faisant nulle mention de la subite transformation de lunettes astronomiques en chaises, Madeleine Malthête-Méliès évoque par contre les intentions des protagonistes, ce qui permet, dans une telle partie introductive, de définir « l'actant-sujet » dans son rapport à l'objet. Exposé oralement et visuellement (sur le tableau noir) par l'orateur diégétique, le projet de se rendre dans la lune est « traduit » par le boniment

304. Cf. Jacques Malthête (dir.), *Essai de reconstitution du catalogue de la Star-Film*, Paris : CNC, 1981, p. 107.

(«atteindre la lune à l'aide d'un obus lancé par un canon géant»), la formulation anticipant les liens intertextuels avec l'œuvre de Verne<sup>305</sup> confirmés ultérieurement par la représentation visuelle (notamment celle de l'obus). Il faut remarquer que l'accompagnement verbal de ce premier «tableau» est particulièrement étoffé par rapport au régime de présence bonimentorielle qui domine dans l'ensemble du *Voyage dans la lune*. En effet, par la suite, la bonimenteuse s'en tient le plus souvent à un titre énoncé au début du plan. Cette intervention bonimentorielle minimaliste correspond à ce que Rick Altman dit des premiers «boniments» de l'ère des nickelodéons: proférés par les projectionnistes eux-mêmes, ils consistaient majoritairement en l'annonce du titre de chaque film<sup>306</sup>. Cette pratique sera ultérieurement remplacée par des plaques de titre montrées grâce à une lanterne, puis par les mentions du générique. Le titrage oral pour lequel opte Madeleine Malthête-Méliès n'est pas sans lien avec l'esthétique des attractions du cinéma des premiers temps, caractérisée par le mode de la juxtaposition. Par contre, le silence de la bonimenteuse est un indice du caractère autosuffisant de la narration visuelle du *Voyage dans la lune*. Par exemple, les plans montrant la chute de l'obus depuis le surplomb lunaire jusqu'aux fonds marins s'enchaînent de façon exceptionnellement rapide, laissant peu de prise au boniment simultané qui se focalise généralement sur les différents lieux montrés. Lorsque l'action domine dans un contexte qui a été explicité verbalement, la bonimenteuse n'intervient plus, laissant à la musique le rôle de souligner rythmiquement les phases du récit. Dans d'autres tableaux, l'absence d'intervention verbale s'explique par la fréquente utilisation de trucages cinématographiques auxquels, comme je l'ai dit, la bonimenteuse ne fait jamais allusion. Ainsi les apparitions symboliques (Phoébé, Saturne, etc.) qui peuplent le ciel lorsque les voyageurs s'endorment ne sont-elles pas commentées, en dépit de la «plus-value culturelle» que ces références pourraient occasionner.

Le titrage des parties débute avec la phrase: «Nous assistons à la construction de l'obus...», prononcée sur un plan dans lequel diverses actions se déroulent qui ne sont pas commentées (par exemple, l'un des futurs explorateurs, poussé par un charpentier

305. Je pense ici à l'édition originale Hetzel du roman *De la terre à la lune* (1865), qui comportait des illustrations.

306. R. Altman, *Silent Film Sound, op. cit.*, p. 140.

qui rabote, tombe dans une cuve). La bonimenteuse utilise à nouveau le « nous », s'associant aux spectateurs dont elle partage concrètement l'emplacement spatial dans la salle. À la fin de la phrase, elle ne baisse pas la voix, comme si d'autres explications allaient suivre; pourtant, elle ne dit plus rien jusqu'au plan suivant, où elle utilise une formule encore plus ramassée que la précédente grâce à la substantivisation (« La fonte du canon géant »). L'intonation, à l'instar des points de suspension dans certains cartons de l'époque muette, donne l'impression que ce sont les images qui prennent le relais du référent verbal. Comme le montrent ces exemples, la bonimenteuse évite systématiquement le recours à des verbes conjugués à l'actif, soit en qualifiant l'action globale par un terme générique (« construction », « fonte »), soit en utilisant une tournure passive (« la rampe de lancement est installée »): son texte relève donc majoritairement du régime « descriptif » au sens large, et met l'accent sur les lieux. Après l'alunissage, elle n'explique pas ce que les personnages sont en train de faire, mais décrit ce qu'ils découvrent (« Au loin, le clair de terre », « À l'intérieur de la lune, la grotte aux champignons géants »). Ces formules elliptiques qui, selon une logique toute « scénique », titrent le tableau en fonction de son décor, ne présentent pas une structure syntaxique destinée à l'organisation du récit, les actions de l'écran conservant leur autonomie. Ce parti pris bonimentoriel convient au *Voyage dans la lune* qui, agencé autour d'un aller-retour encadré par une exposition (discussion des projets et préparatifs) et une conclusion (« Grand défilé, décorations et réjouissances publiques »), et dont le déroulement de la partie centrale s'articule sur une opposition (explorateurs *vs.* Sélénites), offre un fort degré de narrativité. Cette particularité textuelle du *Voyage dans la lune* s'impose, même si l'on sait que Méliès s'est principalement intéressé à composer, comme il le dit dans une lettre rédigée en 1933, « nombre d'originales et amusantes images féériques au-dehors et à l'intérieur de la lune, et à montrer les Monstres, habitants de la lune, en y ajoutant un ou deux effets artistiques »<sup>307</sup>. Les préoccupations majoritairement attractionnelles du « cinéaste » n'interdisent pas que la trame générale de son film obéisse au schéma canonique du récit, dont le boniment peut souligner les

307. Cité par Madeleine Malthête-Méliès, *Méliès l'enchanteur*, Paris: Hachette, 1973, p. 267.

différentes phases. De plus, comme il s'agit de l'œuvre la plus célèbre de Méliès, sa petite-fille a probablement préféré laisser agir la « magie » des images.

Lorsqu'elle bonimente *Barbe-Bleue*, le film le plus long du programme (neuf plans pour une durée totale d'environ neuf minutes), la bonimenteuse se montre par contre considérablement plus loquace. Cette importance de la parole me semble liée à l'origine littéraire du sujet, puisque les tableaux du film illustrent les principaux moments du conte homonyme de Charles Perrault (*La Barbe bleue*, 1697). Cet intertexte s'étend par ailleurs bien au-delà de sa source scripturale, car l'histoire contée par le film convoque également nombre de pratiques qui relèvent, elles, de l'oralité. En 1901, le film de Méliès s'enracine effectivement dans un substrat « culturel » qui bénéficie, depuis près d'un siècle, d'une large diffusion tant en ce qui concerne l'imagerie que le texte, ces deux aspects se trouvant d'ailleurs associés dans les bois gravés à Épinal – précisons que Jean-Charles Pellerin jettera notamment son dévolu sur les contes de Perrault pour donner son essor à sa célèbre industrie d'imagerie, dont les liens avec la transmission orale, toujours actifs au vu de la popularité et de la circulation des sujets, se manifestent ostensiblement dans l'adaptation en images d'airs populaires ou de Cantiques. La pratique de l'illustration légendée n'est sans doute pas sans influence sur la fonction de « titrage » endossée par le boniment du cinéma, lui-même héritier du commentaire des plaques de lanterne magique dont l'esthétique exerça en retour une influence sur l'imagerie du XIX<sup>e</sup> siècle<sup>308</sup>. On peut par exemple supposer que la projection bonimentée d'un film comme la production Pathé *Geneviève de Brabant* (1907) devait rappeler aux spectateurs du début du siècle les spectacles de lanterne magique qui présentaient un sujet identique, à l'instar de la séance domestique évoquée par Marcel Proust (voir I.3). Tom Gunning a d'ailleurs souligné, à propos du film *Geneviève de Brabant*, l'importance du phénomène de répétition (« *retelling* ») qui s'opère lors de l'adaptation filmique d'un conte célèbre<sup>309</sup>. Or *Barbe-bleue* de Méliès s'appuie précisément sur une source de ce type, puisque le conte de Perrault est à compter parmi les motifs

308. Cf. les illustrations in J. Perriault, *op. cit.*, p. 66.

309. Tom Gunning, « Pathé and the Cinematic Tale: Storytelling in Early Cinema », in Michel Marie et Laurent Le Forestier (dir.), *La Firme Pathé Frères 1896-1914*, Paris: AFRHC, 2004, p. 195.

d'origine littéraire qui ont été le plus largement popularisés par l'imagerie xylographique<sup>310</sup>. On peut supposer que les lecteurs de ce genre d'ouvrages réactivèrent certaines habitudes de « narrataires » lors de projections bonimentées aux sujets identiques.

La matière narrative de *Barbe-Bleue* est plus importante que celle du *Voyage dans la lune*, ce dernier n'empruntant aux textes de Jules Verne ou de H.G. Wells que quelques éléments épars. D'ailleurs, lors de la projection de *Barbe-Bleue*, la bonimenteuse commence par expliquer le « pré-texte » du film, c'est-à-dire le passé du seigneur dont les sept premières femmes ont disparu. En outre, ce récit préexistant est lui-même inscrit dans une tradition orale (le conte pour enfants) que le boniment perpétue à sa manière. C'est pourquoi la bonimenteuse peut recourir à des énoncés qu'elle partage avec le public de culture française, renforçant de la sorte la dimension communautaire du « cinéma » parlé. Ainsi, lorsque l'épouse se réfugie sur le toit pour échapper à son mari assassin, Madeleine Malthête-Méliès s'écrie, en se mettant à la place de la victime : « Sœur Anne, ma sœur Anne, ne vois-tu rien venir ? » Puis elle répond en prêtant sa voix à la jeune femme qui, à l'image, est perchée sur la tourelle et scrute l'horizon : « Je ne vois que l'herbe qui verdoie et la route qui poudroie. Ah ! j'aperçois deux cavaliers, ce sont nos frères, vite ! » Assonance et prosodie renvoient à la littérarité du texte original cité. Toutefois, ce surgissement du poétique n'introduit pas une distance entre spectateur et représentation, car ces paroles, quelque étranges qu'elles puissent paraître si on les considère de façon strictement « textuelle », appartiennent à un fonds commun qui permet l'interprétation des actions montrées. Le boniment médiatise le rapport du spectateur au film en faisant de ce dernier un objet connu, familier, certes présenté sous une forme nouvelle qui nécessite le recours à une machinerie, mais rattaché à une expérience de lecteur ou d'auditeur de récits que la plupart des Français possèdent. Faisant l'apologie de la « magie » du cinéma, René Barjavel souligne dans un texte de 1977 le lien qu'il perçoit entre le cinéma de Méliès et l'univers du conte, et cite précisément les vers prononcés ici par Madeleine Malthête-Méliès. Les remarques du romancier me paraissent consonner si profondément avec la problématique des

310. Voir Pierre Louis Duchartre et René Saulnier, *L'Imagerie populaire. Les Images de toutes les provinces françaises du XV<sup>e</sup> siècle au Second Empire. Les Complaintes, contes, chansons, légendes qui ont inspiré les imagiers*, Paris : Librairie de France, 1925, p. 70.

phénomènes intertextuels liés à l'oralité qu'elles méritent d'être mentionnées largement :

«Le cinéma était le grand miracle des temps modernes, plus grand que l'homme sur la Lune et le Concorde. Tous les autres miracles, il les contenait, et il en inventait encore, pour nous les montrer. Il rendait enfin visible le merveilleux qui habitait depuis toujours dans la tête des hommes, installé dès leur enfance par les récits et les contes, puis les livres. Avant d'être imprimées, les histoires avaient franchi les siècles, intactes, par la tradition orale. Quand l'enfant – il les réclamait, il voulait toujours les mêmes – les avait entendues cent fois, il les connaissait par cœur, et il était prêt, à son tour, à devenir conteur.

– Anne, ma sœur Anne, ne vois-tu rien venir ?

– Je ne vois que l'herbe qui verdoie et la route qui poudroie...

J'attendais en frémissant l'arrivée dans le récit de l'immortel et effrayant dialogue. Je ne savais pas exactement ce que signifiaient «verdoie» et «poudroie», mais je voyais ma-sœur-Anne en haut de la tour, regardant vers l'horizon. Et jusqu'au fond du monde il n'y avait que l'herbe et que la route qui verdoie et qui poudroie, c'était terrible, inexplicable, fascinant... Le conte était tout entier en images dans la tête de l'enfant. Chaque mot engendrait son image. Ma-sœur-Anne, pour moi, c'était la sœur aînée d'une copine de la maternelle. [...] Je ne demandais jamais d'explications. Jamais. Je restais avec mes mystères. J'avais la tête pleine de petits nuages blancs. Ça continue... Le cinéma est né du besoin des hommes de voir avec leurs yeux, dehors, devant eux, ces images qu'ils voyaient dans leur tête depuis qu'ils étaient enfants. Le premier metteur en scène, Méliès, a «réalisé» des contes de fées...»<sup>311</sup>

En passant de Perrault à Méliès, Barjavel ne fait pas explicitement référence au boniment des films, mais plutôt à une représentation concrète, dans les féeries filmées, de cette poésie de l'oralité telle qu'elle peut être perçue par un enfant qui visualise mentalement les référents verbaux. La projection bonimentée de *Barbe-Bleue* fait par contre se superposer l'image et la parole, dont les relations réciproques ne s'actualisent pas seulement grâce au «discours

311. René Barjavel, «Miracles à vendre», *Journal du Dimanche*, 4 septembre 1977.



intérieur» de l'auditeur qui connaît le texte de Perrault, mais concrètement à travers l'accompagnement oral. Indépendamment du boniment, il y a néanmoins bien une «voix» sous-jacente aux images muettes que le spectateur peut activer ou non par inférences intertextuelles. Ce processus devait concerner la réception de la majorité des productions fictionnelles de la Star-Film qui, si elles ne se réfèrent pas à des récits préexistants, renvoient néanmoins à une pratique scénique de la prestidigitation qui n'était pas muette. Madeleine Malthête-Méliès n'imitait jamais la voix qui serait celle du magicien sur scène, peut-être parce que les caractéristiques visuelles de l'*adresse* suffisent à la compréhension du film. Dans les performances théâtrales que Méliès « transpose » au cinéma, le verbal a une fonction moins explicative (les trucs demeurent un secret) que *présentative*. La bonimenteuse décide par contre de mettre l'accent sur la voix-narration plutôt que sur la voix-attraction.

Dans *Barbe-Bleue*, la linéarisation induite par le boniment ne se réduit pas à l'ajout d'informations narratives. La parole proférée simultanément aux images tente de coller autant que possible à ce qui est montré afin d'intégrer le spectateur dans la diégèse. Au lieu de commenter les différents tableaux en épousant la position d'une spectatrice avisée, la bonimenteuse de *Barbe-Bleue* se fait littéralement le *porte-parole* des personnages du film. Il en résulte un usage fréquent du discours direct, et de ce fait une forte exploitation du potentiel de la vocalité : les changements de rythme et les exclamations sont beaucoup plus marqués que dans les autres boniments de la séance ; l'intonation y est bien plus vive. Par ailleurs, le discours direct implique le choix d'un « point de vue » dans le récit, notamment lorsque, après le départ en voyage du mari, sa nouvelle épouse brave l'interdiction et pénètre dans la sombre chambre où pendent les cadavres des précédentes conjointes du seigneur. Le boniment nous solidarise en effet avec l'épouse – comme l'image qui, à d'autres moments, nous montre ses rêves et visions – en nous faisant accéder à ses pensées : « J'ouvre un volet... Quelle horreur ! Et la clé, elle est tombée dans le sang ! Cette tache, maudite tache, impossible de l'enlever. » Le pronom personnel « je » ainsi que les accents de panique contribuent à intérioriser l'action, nous incitent à épouser le point de vue de la future victime. Cette subjectivisation s'accorde avec certains éléments visuels, comme les agrandissements par saccades qui affectent la clé, symbole de la transgression de l'ordre du patriarcat.

Cette manière de coller à l'action par le biais de l'imitation d'une source diégétique est également renforcée par la précision dont la bonimenteuse fait preuve au niveau de la synchronisation des paroles avec certains gestes. Il s'agit le plus souvent de mouvements, amples et donc distinctement perceptibles à l'écran, effectués par Georges Méliès lui-même qui interprète le rôle de Barbe-Bleue. Par exemple, après que le père a accepté de donner sa fille à Barbe-Bleue, celui-ci s'active pour signer le contrat de mariage. La bonimenteuse affirme qu'«il faut battre le fer pendant qu'il est encore chaud», alors que l'on voit simultanément Méliès parcourir d'un mouvement horizontal de la main une surface allongée invisible, puis faire mine de la frapper sur le sol. La bonimenteuse remonte ainsi au verbal sous-jacent à la pantomime de l'acteur: elle «traduit» par la parole ce qui avait probablement été transposé en gestes lors du tournage. La synchronisation s'effectue sur une base vocogestuelle lorsque Barbe-Bleue explique à son épouse qu'elle peut se rendre partout dans le château, sauf dans la salle qui leur fait face. À chaque ponctuation de la pantomime – qui recourt à des signes conventionnels clairs – correspond sur l'axe temporel une accentuation verbale ou une pause: Barbe-Bleue tend le trousseau de clés («Je vous confie toutes les clés du château»), parcourt de la main l'espace environnant («Vous pourrez aller partout»), écarte vivement les bras en signe d'exclusion et exhibe une clé gigantesque («sauf...»), pointe la porte de l'index («... cette pièce»). L'expression déictique du discours direct s'ancre totalement dans l'espace diégétique, favorisant l'immersion du spectateur. Cette plongée dans le monde représenté s'effectue également via le «coloris» historique que la bonimenteuse confère à son discours grâce à certains choix lexicaux, comme lorsque le père fait venir le notaire pour rédiger l'acte de mariage en s'écriant: «Holà! tabellion, venez céans avec vos écritoires et vos plumes d'oie!»

À l'instar du boniment des autres films, celui de *Barbe-Bleue* fait totalement l'impasse sur les apparitions oniriques réalisées à l'aide de trucages. Par exemple, Madeleine Malthête-Méliès ne mentionne jamais le diable qui, également interprété par son aïeul, apparaît à plusieurs reprises, sortant d'abord avec force fumigènes d'un livre ouvert. L'aspect féerique demeure du seul ressort de l'image. Une unique mention orale renvoie indirectement à un «truc»: lorsque des femmes apparaissent en surimpression, le commentateur précise que «les sept femmes de Barbe-Bleue ressusci-

tent». Nous sommes toutefois à la toute fin du film, là où s'opère la conclusion des enjeux narratifs («Tout est bien qui finit bien», dit la bonimenteuse): cette apothéose théâtrale n'appartient déjà plus tout à fait au domaine du récit, ainsi que le note Elisabeth Ezra, qui, dans sa perspective metzienne, considère ce type de *finale* (et celui de *Barbe-Bleue* en particulier) comme l'une des modalités d'«inserts non diégétiques»<sup>312</sup>.

Le boniment de Madeleine Malthête-Méliès, nécessairement soumis au contexte commémoratif dans lequel il prend place, ne permet bien sûr d'illustrer qu'un nombre restreint d'options définissant les rapports qu'un locuteur intégré à un dispositif audiovisuel instaure avec les images, le spectateur et la machinerie. On constate néanmoins que le boniment simultané aux images met l'accent sur le représenté, alors que la mécanique (du dispositif) ou la technicité (des trucages) est occultée. En dépit de son statut de reconstitution, cette séance bonimentée témoigne non seulement de l'intérêt rencontré par le «cinéma» parlé à la fin du XX<sup>e</sup> siècle – fût-il lié à un «auteur» – mais aussi de diverses stratégies dont dispose (et disposait) le bonimenteur de vues animées.

312. Elisabeth Ezra, *Georges Méliès. The Birth of the Auteur*, Manchester; New York: Manchester University Press, 2000, p. 38.



### III. LE « BONIMENT » OVER DU ROMAN D'UN TRICHEUR

#### III.1 UN USAGE SINGULIER DE LA VOIX

Sorti en 1936, *Le Roman d'un tricheur* de Sacha Guitry est un film tout à fait symptomatique de l'absence de standardisation qui caractérise la période de l'« interrègne », dont cette œuvre représente en quelque sorte le chant du cygne. Cette même année, Charlie Chaplin, figure emblématique de la résistance à l'institutionnalisation du cinéma parlant, réalise le dernier de ses films qui échappent au « 100% talkie », *Les Temps modernes* (*Modern Times*, 1936). La fin des années 30 verra la généralisation d'une intégration normalisée de la voix au film, majoritairement cantonnée au synchronisme vocolabial strict ou à une forme « classique » de voix-over qui, dans le cinéma hollywoodien, connaît un véritable essor à partir de 1939<sup>313</sup>. Or l'usage de la voix-over dans le film de Guitry se distingue radicalement de ces pratiques qui sont en passe de devenir dominantes, quand bien même le cinéaste français est fréquemment considéré comme « l'inventeur » du procédé<sup>314</sup>. *Le Roman d'un tricheur* établit plutôt une filiation entre le boniment du « cinéma » parlé et les voix non synchrones (mais synchronisées) du parlant. En dehors de certaines productions de la Nouvelle Vague française réalisées une trentaine d'années plus tard – et l'on sait l'intérêt que portaient Truffaut et Resnais au film de Guitry, on ne retrouvera guère une telle liberté dans l'utilisation de la voix-over

313. Voir S. Kozloff, *Invisible Storytellers...*, *op. cit.*, p. 142. Cette standardisation n'interdit pas que certains films hollywoodiens des années 1940-1950 offrent, grâce aux manifestations *over* qu'ils proposent, une organisation énonciative complexe qui n'obéit pas aux canons du récit classique (*Laura, Le Secret derrière la porte*, les films de Mankiewicz, ...).

314. Lorsqu'il commente les « films noirs » hollywoodiens récemment distribués en France, Nino Frank dit par exemple à propos de l'intervention d'un « narrateur ou commentateur » : « Sacha Guitry a – le premier – utilisé ce procédé dans *Le Roman d'un tricheur* (Nino Frank, « Un nouveau genre policier. L'aventure criminelle », *L'Écran français*, N° 61, 28 août 1946, p. 14).

que dans des pratiques marginales, par exemple dans le cinéma d'animation<sup>315</sup> ou dans des films où la voix est considérée comme un ajout ultérieur, et non comme une donnée consubstantielle de l'œuvre originale. Il en va par exemple ainsi des versions à voix-over de *La Ruée vers l'or* (*Gold Rush*, Charlie Chaplin, 1925), qui connaît en 1942 une ressortie comprenant un commentaire rédigé et proféré par le cinéaste lui-même, ou de la «voix-over de traduction» de *Man about Town*, version du *Silence est d'or* de René Clair (1947) destinée à un public anglophone (Maurice Chevalier y intervient en tant que narrateur pour conter les principaux événements en anglais, les acteurs visualisés n'étant pas doublés). René Clair, dont la réflexion sur le parlant au moment de sa généralisation et les tentatives d'exploiter dans sa pratique de cinéaste tout le potentiel du cinéma sonore ont marqué le milieu cinématographique de l'époque, déclarait à la fin des années 50, à propos du projet qu'il avait nourri une trentaine d'années auparavant (notamment dans *Le Million*) d'introduire au cinéma le «monologue intérieur» :

«Quand, plus tard, apparut *Le Roman d'un tricheur*, je compris que j'avais eu tort de ne pas persévérer. Sacha Guitry avait inventé de son côté le principe du commentaire sans avoir jamais entendu parler de ma tentative et, en quelque sorte, par hasard. [...] Paradoxalement, c'est peut-être par une sorte de méconnaissance du cinéma que Sacha Guitry fit son chef-d'œuvre cinématographique.»<sup>316</sup>

On constate qu'en dépit de sa volonté de diminuer les mérites d'un cinéaste qu'il associe encore, probablement avec quelques ressentiments, à un discours de valorisation du théâtre par rapport au cinéma, René Clair n'hésite pas à souligner le caractère novateur du film de Guitry. Il paraît néanmoins difficile de qualifier d'«invention» la réactualisation d'un type de rapport entre l'image et le verbe qui, somme toute, résulte simplement de la fixation phonographique de ce boniment auquel René Clair avait pu assister au

315. À l'exemple du *Brave soldat Svejk* de Jiri Trnka (1955), dans lequel les logorrhées interminables du personnage créé par Hasek, représenté sous la forme d'une poupée à la bouche immobile et au sourire perpétuel, sont proférées par un locuteur qui prête par ailleurs sa voix à tous les autres personnages.

316. René Clair, *Comédies et commentaires*, Paris: Gallimard, 1959, p. 91.

début des années 10, et qu'il met en scène dans la séquence inaugurale du *Silence est d'or*. Certains critiques contemporains du *Roman d'un tricheur* n'ont d'ailleurs pas manqué de souligner combien la voix-over du film de Guitry s'inscrivait dans la filiation de pratiques antérieures: le chroniqueur de la revue *Caméra* écrit par exemple que le «film [est] muet car les acteurs se taisent, parlant peu car M. Sacha Guitry discourt du commencement à la fin... La formule n'est pas neuve: il raconte un film muet», tandis que Bardèche se demande dans *L'Assaut* si «ce commentaire continu» n'est pas «autre chose que l'ancien sous-titre»<sup>317</sup>.

Tiraillé entre l'appropriation d'une nouveauté technologique et la reconstitution de pratiques scéniques qui engagent la performance orale, *Le Roman d'un tricheur* constitue un objet d'étude approprié à la perspective adoptée dans le présent ouvrage. Il nous permettra non seulement de réfléchir plus avant sur les liens entre boniment et voix-over, mais aussi d'amorcer une discussion sur l'opposition entre voix-narration et voix-attraction, deux tendances que ce film présente tour à tour en conciliant une organisation narrative attribuée à une instance *over* totalement intégrée au texte filmique avec une exhibition de la vocalité sur un mode qui confine parfois à l'adresse au spectateur.

Si la façon dont Guitry use de la voix-over dans *Le Roman d'un tricheur* est originale, c'est parce qu'elle accompagne des images muettes alors que les acteurs visualisés remuent les lèvres. Grâce à l'éviction permanente de la voix synchrone – parti pris radical qui n'a pas été sans insupporter à l'époque le critique Claude Aveline, qui aurait souhaité que Guitry usât du procédé «avec la plus extrême mesure, en quelques endroits seulement»<sup>318</sup> – *Le Roman d'un tricheur* se distingue de la très grande majorité des films en voix-over qui s'en tiennent à une suppression ponctuelle des voix *in* (voir par exemple le prologue d'*Ève* *All about Eve*, Joseph L. Mankiewicz, 1950). Dans un ouvrage où il ne prend en compte que les films projetés dans les salles avant 1939, Jean Keim souligne l'existence d'une filiation entre l'époque «muette» et les débuts du parlant en rapprochant le bonimenteur – notons qu'il recourt précisément à ce terme – de l'instance *over* qui narre oralement le *Roman d'un tricheur*.

317. Cité par Dominique Desanti, *Sacha Guitry. 50 ans de spectacle*, Paris: Grasset; Fasquelle, 1982, p. 206.

318. Claude Aveline, «*Le Roman d'un tricheur – Jenny*», *La Revue bleue*, 17 octobre 1936, repris in *Chroniques d'un cinéphile*, Paris: Séguier, 1994, pp. 199-200.

«On pouvait remplacer par des paroles enregistrées le bonimenteur des premiers films sans titres de l'époque héroïque, qui, debout devant l'écran, expliquait aux spectateurs l'action [...]. Cette méthode d'ailleurs n'a été employée que pour les documentaires et les actualités. En ce qui concerne les grands films sonores, on ne peut guère citer de cette catégorie que *Le Roman d'un tricheur*, de Sacha Guitry, où l'auteur-acteur raconte de sa voix bien connue l'histoire qui se déroule sur l'écran, interprétée par des artistes aphones.»<sup>319</sup>

Keim considère *Le Roman d'un tricheur* comme le seul film de fiction jugé important qui ait entrepris de restituer par des sons fixés la voix vive du «muet». Il est vrai que plus que tout autre, le film de Guitry présente une esthétique qui s'origine dans le «cinéma» parlé, alors que les productions hollywoodiennes contemporaines, dont la très grande majorité est dépourvue de voix-over, subissent plutôt l'influence de la radio et de l'industrie discographique. Si les artistes sont «aphones», c'est que l'énonciateur *over* du *Roman d'un tricheur* «prend la parole» à ses personnages ou la leur donne alors qu'ils n'articulent visiblement aucun son. Ce type de rapport entre l'image et la voix exhibe l'omnipotence du verbal sur la représentation. Située dans une zone intermédiaire entre le doublage et la «voix-over de traduction» telle qu'on la trouve encore aujourd'hui dans les salles et sur le petit écran de certains pays d'Europe de l'Est (Russie, Pologne)<sup>320</sup>, la parole du cinéaste se permet des va-et-vient incessants entre l'identification aux personnages et la distance d'un commentateur amusé. D'ailleurs, les langues étrangères sont évincées du film, le narrateur *over* se contentant par exemple de dire à propos des terroristes qu'ils «échangèrent quelques mots à voix basse en russe». Dans la séquence du joaillier, Guitry s'exprime à la place de la voleuse comme si celle-ci rendait compte verbalement de ses

319. J. Keim, *op. cit.*, p. 41.

320. Voir par exemple le témoignage de Pascal Mériegeau, qui écrit, après avoir assisté à des projections de films français à Moscou: «En ouverture du Festival du Cinéma français, *Un long dimanche de fiançailles*. La projection commence. Silence religieux. Une voix masculine russe dit les dialogues des hommes, une voix féminine ceux des femmes, on entend également la voix des acteurs, le principe du *voice over* produit toujours des effets inattendus. Les Russes y sont habitués» (*Le Nouvel Observateur*, N° 2113, 5 mai 2005). Cette fonction de traduction héritée du boniment est parfois intégrée aux films, comme dans *Berlin Express* (Jacques Tourneur, 1949), film dont la voix-over traduit ou résume pour le public américain certaines phrases prononcées en français lorsque les actions diégétiques sont situées à Paris.



moindres gestes. Cette posture descriptive aux accents railleurs met en évidence l'extériorité de l'énonciateur *over* par rapport à l'image. Ostensiblement, le locuteur se situe en retrait, mais décrit l'action dans ses moindres détails en se souciant de la synchronisation. Bien qu'il restitue une situation dialogique qui postulerait d'«incarner» le personnage parlant, Guitry s'adonne dans de tels cas à une technique que réprouvait avec force dans un texte de 1911 le bonimenteur Georges Dalbe, dont la réaction atteste de la fréquence de cette façon de faire chez ses confrères: «Règle générale, l'on doit employer le dialogue, il faut se défier des: 'Il fait ceci... il s'en va', car tout le monde s'aperçoit des sorties: le cinématographe n'ayant jamais eu pour but d'intéresser les aveugles.»<sup>321</sup> Chez Guitry, le peu de pertinence des informations convoyées verbalement par rapport à ce que l'on peut nettement distinguer à l'écran fonctionne au bénéfice des fonctions phatique et conative, le référent s'avérant finalement bien moins important que le ton employé. L'ironie et le mordant participent de cette distance quasi physique avec la représentation et, comme toute forme d'humour (d'autant plus lorsque les «mots d'auteur» y sont si nombreux), instaurent une connivence avec le public. Le recul dont Guitry témoigne constamment en tant que locuteur par rapport aux énoncés qu'il profère passe essentiellement par les modalités de l'actualisation orale de son texte écrit, dans la mesure où il use fréquemment d'«ajustements» au sens de Goffman, c'est-à-dire d'«indices vocaux pour s'assurer que les limites et le caractère du segment censé être une citation se démarquent bien du flux normal»<sup>322</sup>. Guitry se plaît en effet à imiter les voix des personnages lorsqu'il énonce des répliques à leur place: il reproduit le cri du sourd-muet agonisant, imite l'accent du terroriste russe ou celui des habitants de Clermont-Ferrand lorsque son personnage se donne une fausse identité, change sensiblement de ton lorsqu'il parle à la place d'une femme, etc. Ce travail sur les voix est une composante de l'atmosphère drolatique et désinvolte de ce film où tout repose sur le *jeu*, dans les différents sens du terme, c'est-à-dire, chez Guitry, sur les déploiements ludiques du *je*. Philippe Le Guay a souligné combien la jubilation nombriliste de l'auteur passe également par la voix: «Pour Guitry, l'Autre est un je. Tous les

321. Georges Dalbe, *art. cit.*, p. 32.

322. Erwin Goffman, *Façons de parler*, Paris: Minit, 1987 [1<sup>re</sup> édition 1981], p. 182.

masques renvoient à l'identité monolithique de la première personne, au narcissisme, au miroir... à la voix.»<sup>323</sup> La multiplication des déguisements dont use le tricheur participe de la même logique: le personnage-narrateur acquiert le pouvoir d'incarner tous les protagonistes du film.

En dépit de l'ancrage diégétique de l'énonciateur verbal dans une situation-cadre sur laquelle le film revient par intermittence (le récit enchâssé se présentant comme une visualisation du journal que le personnage est en train d'écrire à la terrasse d'un café), la voix semble provenir d'un ailleurs. Nous n'entrons pas «dans la tête» de l'écrivain pour assister à la visualisation d'un monologue intérieur – procédé où d'ailleurs l'image tend traditionnellement à supplanter progressivement le verbal, alors qu'ici la voix ne s'interrompt presque jamais – mais nous écoutons le locuteur *nous parler du film* qui est projeté. En effet, les références du verbal ne portent pas seulement sur le contexte ou le contenu narratif des actions montrées, mais sur certains détails de leur (re)présentation à l'écran. La précision de la synchronisation, manifeste à tous les niveaux (vocolabial, vocogestuel, vococinétique), produit l'effet inverse de celui visé par un usage plus conventionnel de la voix synchrone: le travail même de synchronisation humaine prend ici le pas sur le synchronisme. Lorsque Truffaut dit du film que «son immense mérite est d'avoir fait oublier cette particularité [la présence continue d'une voix-over] au point que des spectateurs interrogés à la sortie de la salle croiront avoir vu un film joué et parlé directement»<sup>324</sup>, il néglige l'exhibition volontaire dont cette voix participe et valorise, sur la base d'une évaluation fort hypothétique de sa réception, une «transparence» qui n'est pas la seule façon de s'assurer la participation du public. Ce que Truffaut semble vouloir souligner, c'est que la voix-over du *Roman d'un tricheur*, instance de médiation entre le spectateur et le monde des images, n'a pas occasionné de rejet de la part du public. La complicité qu'elle instaure avec ce dernier me semble en effet garantir une lecture régie par une posture d'*adhésion* au film.

Si le narrateur semble regarder les images avec nous, il affiche néanmoins son emprise sur le sens des images et sur le film lui-même. La séquence où il se plaît à imaginer la garde de Monaco

323. Philippe Le Guay, «Le double boîtier», *Cinématographe*, N° 86, 1983, p. 17.

324. François Truffaut, «Préface», in *Sacha Guitry. Le Cinéma et moi*, Paris: Ramsay, 1977, p. 19.

effectuant autant de pas en arrière qu'elle en a effectués en avant est la démonstration la plus évidente de ce pouvoir de manipulation conféré à la voix : après un point de synchronisation de mise en chaîne (au niveau photogrammatique) fortement exhibé (le narrateur s'écrie « tout à coup » en élevant la voix, alors qu'à cet instant précis l'image s'immobilise), c'est le sens même du défilement des images qui se soumet aux desiderata ludiques de l'énonciateur verbal. Or, rappelons-le, ce « trucage » réalisé à la projection avait cours à l'époque du cinéma des premiers temps. Cette mise en évidence du matériau filmique peut donc également se comprendre comme une réminiscence des séances où le bonimenteur créait explicitement une médiation entre les spectateurs et le projectionniste.

Le rapprochement avec le « cinéma » parlé s'impose également du fait que les sons dont l'origine est explicitement située dans l'action sont très rares et ponctuels, à l'instar de ce que l'on pouvait obtenir avec des effets sonores réalisés dans la salle à l'époque du muet. Il en va par exemple ainsi des sonneries répétées des clients de l'hôtel qui appellent le jeune liftier disparu dans la chambre de la comtesse. Dans ce cas toutefois, leur mécontentement s'exprime également par un brouhaha indistinct et croissant : ce type d'occurrence vocale collective à synchronisation large, caractéristique de l'« ambiance réaliste » rudimentaire des tout premiers films parlants (ou de ceux qui furent sonorisés ultérieurement), apparaît à plusieurs reprises dans *Le Roman d'un tricheur* (cris et applaudissements des soldats dans la taverne, manifestations d'enthousiasme des joueurs de roulette lorsque le zéro sort plusieurs fois consécutivement, etc.). Par ailleurs, la musique, qui n'a le plus souvent pas à composer avec d'autres sons diégétiques, rappelle à de nombreux égards l'accompagnement musical du « muet » en procédant à de brèves citations ou à de brusques changements de mélodie afin de coller à l'atmosphère décrite, notamment lorsqu'il s'agit d'imiter une source diégétique. De façon générale, la musique du *Roman d'un tricheur* est soumise au contenu des images, et non à la voix-over qui court parallèlement. C'est surtout dans les (rares) moments où la voix-over se tait complètement que la musique du film s'apparente le plus à l'accompagnement des projections « muettes », par exemple durant l'arrestation des terroristes russes, où chaque action est ponctuée musicalement, mais n'occasionne aucun son diégétique. Cette réduction extrême de l'éventail des occurrences sonores s'explique certes par les limites de la technologie disponible en

France au milieu des années 30, les ingénieurs du son ne disposant à cette époque que d'une étroite bande passante. Toutefois, Guitry fait de cette contrainte une composante stylistique essentielle de son film, prenant à la lettre cette nécessité qui, selon Michel Chion, vaut pour les productions des toutes premières années du parlant : « Quand la voix parlait devaient s'effacer les rumeurs et les bruits. »<sup>325</sup> Dans la scène où la voleuse s'avise de pratiquer un trou dans la paroi de la chambre d'hôtel, la musique provenant d'un gramophone qu'elle a préalablement actionné pour couvrir le bruit de la perceuse ne semble pas provenir de la diégèse, mais évoque également la « musique d'écran » qui s'ancrait par moments dans une source visible potentielle. En effet, le bruit de la chignole est non seulement étouffé comme l'exige le contexte narratif, mais totalement inaudible (ce qui est peu vraisemblable selon une stricte logique diégétique), la musique planant sur l'image en créant la même homogénéité sonore que la voix-over. En ce qui concerne la sélection des sons, il est à noter que Guitry privilégie dans cette séquence (comme dans l'ensemble du film) la composante humaine, puisque ni les caractéristiques phonographiques de la restitution de la musique, ni le fracas de la machine-outil ne se font entendre.

### III.2 GUITRY ET L'ORALITE : D'UNE CERTAINE « MODERNITE » CINEMATOGRAPHIQUE

Ces remarques liminaires témoignent du statut paradoxal du *Roman d'un tricheur* : il s'agit d'une œuvre à la fois volubile et proche du cinéma muet. En fait, en dépit du primat qu'il accorde au verbe, ce film s'inscrit dans les expérimentations au caractère hybride des premières années du parlant, à l'époque où cinéastes et critiques prônaient le modèle du cinéma « sonore ». Dans un texte contemporain de la sortie du film de Guitry, Erwin Panofsky considère d'ailleurs *Le Roman d'un tricheur* comme une entreprise nostalgique visant à « combiner les vertus du son et de la parole avec celles du jeu du muet »<sup>326</sup>.

La parenté entre *Le Roman d'un tricheur* et le spectacle bonimenté semble liée au parcours même de Sacha Guitry qui, avant

325. Michel Chion, *La Toile trouée*, Paris : Éditions des Cahiers du cinéma, 1988, p. 28.

326. Erwin Panofsky, « Style et matière du septième art » [1936], in *Trois essais sur le style*, Paris : Gallimard, 1996, p. 134.

de revenir au cinéma en 1935 avec *Pasteur*, avait l'expérience d'un seul film, *Ceux de chez nous*, réalisé en 1914-1915 et dont il accompagnait la projection d'une conférence donnée avec son épouse. Ainsi Guitry fut-il « bonimenteur » avant même d'avoir véritablement commencé sa carrière de cinéaste, une profession dont il ne pouvait sans doute pas concevoir l'intérêt pour autant que l'œuvre réalisée restât muette. Alex Madis, journaliste et fervent admirateur de Guitry, mentionne dans sa biographie consacrée au « maître » en tant qu'auteur dramatique un témoignage qui, bien qu'ayant la fiabilité toute relative d'un lointain souvenir (Madis rédige son texte trente-cinq ans plus tard), permet de supputer quels furent les effets de cette projection bonimentée du milieu des années 10 sur le travail ultérieur de Guitry :

« Mais tout d'un coup, on ne les voyait pas seulement, on les entendait. Postés au proscénium, au-dessous de l'écran, M<sup>me</sup> Charlotte Lysès et Sacha exprimaient à haute voix leur admiration pour « ceux de chez nous » et leurs mouvements de lèvres en image correspondaient avec une telle exactitude aux mots prononcés par eux, en chair et en os, que l'illusion était totale. Le film parlant était né ! »<sup>327</sup>

La description que fait Madis de cette performance vivante fait écho à l'usage de la voix-over dans *Le Roman d'un tricheur*, où le locuteur Sacha Guitry prête sa voix à d'autres. La vision téléologique du biographe selon laquelle la projection de *Ceux de chez nous* peut être considérée comme « l'avènement » du cinéma parlant l'amène à souligner une perfection illusionniste du synchronisme que l'on retrouve également dans le film de 1936. Les modalités spectaculaires de cette projection sont toutefois fondamentalement différentes de celle du dispositif du cinéma parlant : exhibés sur le proscénium, les locuteurs sont extérieurs à la représentation écranique. Le synchronisme n'est pas naturalisé, mais constitue l'une des attractions de la séance. La parole se voulait instrument de conviction, puisqu'il s'agissait pour Guitry de faire reconnaître la grandeur des personnalités de France montrées à l'écran (Edmond Rostand, Claude Monet, Anatole France, Lucien Guitry, etc.). Cette visée patriotique conduisit vraisemblablement à une exploitation optimale des circonstances de la performance

327. Alex Madis, *Sacha*, Paris : Éditions de l'Élan, 1950, p. 266.

orale, la rhétorique du verbe servant l'éloge. D'ailleurs, lorsque Sacha Guitry réitère l'expérience d'une conférence accompagnant une projection avec *De 1429 à 1942, ou De Jeanne d'Arc à Philippe Pétain*, le 6 mars 1942, au Théâtre de la Madeleine<sup>328</sup>, il entend également faire l'apologie d'importantes personnalités nationales issues de la politique et des arts. Les conditions de cette projection bonimentée unique, occultée dans la majorité des filmographies (peut-être en raison du caractère éphémère de cette performance, ou parce qu'elle s'inscrit dans une période peu glorieuse de la carrière du « maître » sur laquelle personne ne s'attarde), furent tout à fait particulières, la radicalité de la démarche étant révélatrice de la place accordée au texte et à la voix dans le cinéma de Guitry. Avant que le recueil de textes rédigés par Guitry et d'autres écrivains (Duhamel, Cocteau, Morand, Giraudoux, etc.) intitulé *De Jeanne d'Arc à Philippe Pétain* ne soit publié en 1944 dans une édition luxueuse au tirage limité, Guitry organise une « lecture » de ces textes. Or le film qui défile simultanément à son discours consiste basiquement à montrer les pages mêmes de l'ouvrage lues par le conférencier. L'écrit présent dans *Le Roman d'un tricheur* sous la forme du journal intime, relais provisoire vers le *flash-back* en voix-over, s'impose ici continuellement, affichant la matérialité du livre et la dimension graphique du lettrage. Pour Guitry, la voix est avant tout la manifestation d'un texte sous-jacent qui est « l'œuvre de l'auteur », c'est pourquoi il incarne souvent lui-même les personnages qui déclament les répliques principales. Comme le note Alain Keit, la monstration d'un livre ou de l'écriture joue fréquemment dans les films de Guitry un rôle de transition vers l'oralité, notamment sous la forme de la voix-over<sup>329</sup>. Guitry semble se délecter de cette animation que la voix d'un acteur offre au support figé qu'est le livre. Dans *Le Roman d'un tricheur*, la fonction de relais endossée par l'écrit apparaît d'ailleurs à deux niveaux puisque non seulement le récit enchâssé est présenté comme la visualisation du texte des mémoires, mais aussi parce que le récit métadiégétique de l'arrestation des terroristes russes illustre ce que le narrateur dit avoir lu à l'époque dans un quotidien (« et je lus qu'il était arrivé ceci : ... »). Le fait que, dans ce second cas, aucune voix n'accompagne la représentation visuelle de l'arrestation est lié

328. Voir D. Desanti, *op. cit.*, pp. 361-362.

329. Alain Keit, *Le Cinéma de Sacha Guitry. Vérités, représentations, simulacres*, Liège: Céfal, 1999, p. 112.

au statut secondaire de la source textuelle supposée: comme le texte de la presse n'est pas censé avoir le même ton que le récit du tricheur, ce personnage disparaît en tant qu'origine verbale.

L'image du film *De Jeanne d'Arc à Philippe Pétain* n'a pas à «illustrer» le référent du texte oral, elle est le texte, s'y réfère comme à un objet du monde. Dans un tel cas, la notion de «redondance» perd plus que jamais sa pertinence puisque le référent iconique donne lui-même lieu à un référent verbal. La diégèse filmique est réduite, en termes de construction de monde, à «l'ameublement» minimal qu'est la présence du livre. En outre, le spectateur du film est appelé à effectuer la même opération de lecture que le conférencier, ce qui souligne combien ce dernier est lié, au sein du dispositif, à l'espace concret de la réception dont il assure la médiation avec la représentation visuelle. Le livre dont le film montre les pages ne contient toutefois pas que du texte, mais également de nombreuses illustrations (photographies, peintures, dessins)<sup>330</sup>: la fonction endossée par le verbal dans son rapport à l'image est ainsi dédoublée lors de la projection du film, le boniment commentant une image filmique qui comprend elle-même des images assorties d'un texte. Sacha Guitry aime s'appropriier par la parole une iconographie préexistante. Dans *Le Roman d'un tricheur*, il se sert d'images d'archives dont l'intégration donne par moments l'impression que le film relève du documentaire (par exemple dans la séquence où il présente la ville de Monaco). Cette démarche incite le spectateur à opérer une mise en relation explicite du texte et de l'image. Elle concerne également l'auditeur d'une émission radiophonique comme *Cent merveilles* réalisée par Guitry, dont la diffusion s'accompagna de l'édition d'un ouvrage qui permettait au spectateur «de se reporter à la page indiquée par Sacha Guitry et d'avoir sous les yeux l'objet de l'émission»<sup>331</sup>.

Lorsque le moyen métrage *De Jeanne d'Arc à Philippe Pétain* fut projeté dans le cadre de la rétrospective consacrée à Guitry au Festival de Locarno de 1993, Yann Lardeau, dans un texte du catalogue édité à cette occasion, émet l'hypothèse selon laquelle ce film aurait été accompagné d'un commentaire verbal. Sa description

330. Voir Sacha Guitry, *Cinquante ans d'occupations*, Paris: Presses de la Cité, 1993 [1<sup>re</sup> édition 1947], pp. 1033-1104.

331. *Ibid.*, p. 1105.

montre qu'à l'instar du *Roman d'un tricheur*, ce film comprend d'autres sons que la seule voix du conférencier :

« Il s'agit d'un « plan » unique montrant les pages de l'exemplaire dédicacé au maréchal Pétain tournées les unes après les autres. Cinquante-huit minutes de banc-titre et de silence entrecoupé de temps à autre par la lecture d'un poème ou le fragment d'une composition musicale, un air d'opéra. Comment se plaçait le texte de Guitry et ceux des autres écrivains sollicités [...] par rapport aux images ? Pourquoi ces trous si énormes dans la bande-son que, musique ou mots, ce sont les sons qui apparaissent comme des accidents ? En fait, il semble bien que Sacha Guitry ait conçu *De Jeanne d'Arc à Philippe Pétain* selon le même modèle que la version initiale de *Ceux de chez nous*. »<sup>332</sup>

La projection a dû voir se côtoyer un ensemble de voix-over fixées et synchronisées mécaniquement et la voix du conférencier. On trouve ce principe des « vides » laissés dans la bande-son dans les premiers films sonores nippons (par exemple lorsque la musique extradiégétique s'interrompt dans *Une auberge à Tokyo*, Yasujiro Ozu, 1935) qui étaient prévus pour connaître une projection accompagnée par un *benshi*. Ce principe de soustraction se perpétue aujourd'hui avec les pistes-son des vidéoclips prévus pour les jeux de karaoké. La fonction de médiation d'une instance vocale extérieure à l'univers diégétique qui caractérise les deux « conférences » données par Guitry me semble également s'appliquer au *Roman d'un tricheur*, dans lequel la voix-over se charge de la conduite du récit, s'abandonne à des digressions sous forme de commentaires explicatifs et d'argumentations, ou énonce des répliques attribuées aux personnages visualisés. Le rapport particulier que Guitry instaure ici avec l'audiospectateur par l'intermédiaire de la voix s'applique plus généralement à l'ensemble de son cinéma, à propos duquel Dominique Païni note qu'il s'agit d'un « dispositif très étrange, qui fait alterner l'adresse au spectateur et le monologue », permettant en quelque sorte au locuteur interne au discours filmique de « franchir l'écran et de descendre dans la salle »<sup>333</sup>.

332. Yann Lardeau, « Les pages arrachées au Livre de Satan », in *Sacha Guitry, cinéaste*, Locarno ; Crisnée : Éditions du Festival de Locarno ; Yellow Now, 1993, p. 45.

333. Vincent Amiel et Noël Herpe, « C'est la parole qui incarne. Entretien avec Dominique Païni », *Double Jeu*, N° 3 (« Sacha Guitry et les acteurs »), 2006, p. 146.



*Le Roman d'un tricheur* fait montre d'une diversité au niveau des rapports entre la voix et l'image qui ranime l'hétérogénéité des pratiques du « cinéma » parlé. *Ceux de chez nous* a d'ailleurs connu, en ce qui concerne sa dimension sonore, deux autres traitements : en 1939, Guitry enregistre un commentaire *over* ; en 1952, il monte une version allongée dans laquelle il apparaît lui-même à son bureau, lisant le commentaire comme le personnage du *Roman d'un tricheur* écrit, attablé, son histoire. La suppression d'un régime vocal mixte à la fin des années 30 correspond au processus d'institutionnalisation qui est contemporain de ces deux versions. En 1952, par contre, Guitry retravaille son film pour la télévision, c'est-à-dire qu'il l'adapte à un espace de diffusion qui redonne une nouvelle dimension à l'oralité : il peut alors se permettre un mode de communication plus présentationnel que « transparent ». Selon Bernard Eisenschitz, cette version parlante constitue une sorte de recréation de la projection bonimentée des années 10 :

« Il offre même, à ce spectateur, un exemple de l'interaction entre le cinématographe et la représentation *live*, donnant à la fois une idée de ce que pouvait être sa causerie accompagnée de projections, ou la projection accompagnée de ses commentaires. »<sup>334</sup>

Sacha Guitry semble donc être l'un des principaux représentants du transfert de pratiques orales vivantes vers leur enregistrement phono-cinématographique. D'ailleurs, lorsqu'un cinéaste également féru de théâtre et d'interventions radiophoniques comme Orson Welles utilise sa propre voix de façon peu orthodoxe (par exemple au début de *La Splendeur des Amberson*, où le narrateur extradiégétique converse avec « ses » personnages qui lui répondent comme s'il appartenait au même monde qu'eux) pour accompagner *over* les images d'un film de fiction, Bettina Knapp décèle chez lui une influence de ce qu'elle appelle le « cinéma commenté » qu'aurait inauguré *Le Roman d'un tricheur*<sup>335</sup>.

En raison de sa passion pour le théâtre, Sacha Guitry fut donc à plusieurs reprises amené à placer la projection cinématographique sous le signe de l'oralité en considérant la séance comme une performance unique. Le caractère labile de telles pratiques explique

334. Bernard Eisenschitz, « Ceux de chez nous », *Sacha Guitry, cinéaste, op. cit.*, pp. 151-152.

335. Cité in Chuck Berg et Tom Erskine, *The Encyclopedia of Orson Welles*, New York : Facts on File, 2003, p. 132.

bien sûr qu'il n'en demeure souvent pas trace dans les écrits sur le travail du cinéaste. Les exemples de *Ceux de chez nous* et *De Jeanne d'Arc à Philippe Pétain* ne constituent toutefois pas des cas isolés dans la carrière de Guitry-cinéaste, ainsi que l'indique Claude Arnaud en citant deux autres manifestations vocales *live* intégrées à une séance de cinéma :

«Le 28 mars 1942, il présente au Gaumont-Palace un court-métrage (*La loi du 21 juin 1907*). Gertrude (Arletty) rentre chez son père: elle ne pourra pas épouser Gaston (Fernand Gravey). Leur dernière chance: consulter un avocat. Dans la salle du Gaumont, quelqu'un se lève (Fernand Ledoux), monte sur scène, et se propose de les aider. Entre l'écran et la scène commence un étrange dialogue [...]. Déjà, en 1938, il [Guitry] avait interrompu la projection de *Quadrille* pour jouer avec Gaby Morlay le troisième acte de la pièce, avant de laisser à la pellicule le soin de conclure.»<sup>336</sup>

De telles interventions d'un locuteur présent dans la salle n'est pas sans évoquer un type de spectacle hétérogène comme le music-hall, que j'associerai à l'usage de la voix-attraction (voir chapitre V). *Le Roman d'un tricheur* est indubitablement le film de Guitry qui ressortit le plus ostensiblement à une pratique de l'oralité. C'est pourquoi il sera rapproché d'une forme de «modernité» cinématographique à l'époque de la Nouvelle Vague, c'est-à-dire à un moment de l'histoire du cinéma qui, selon Noël Burch, connaît un retour aux pratiques sonores de «l'interrègne»<sup>337</sup>. D'ailleurs certains représentants de la Nouvelle Vague se réclameront de Guitry, notamment pour son travail sur la voix. On sait en effet l'éloge que François Truffaut – un cinéaste coutumier des voix-over, qu'il utilise notamment dans *Jules et Jim* (1962) et *Les Deux Anglaises et le continent* (1971) – fait à Guitry en 1977 dans la préface au recueil d'articles écrits par ce dernier sur le cinéma<sup>338</sup>. À la sortie d'*Hiroshima mon amour*, Alain Resnais revendique quant à lui une filiation avec ce film en affirmant: «Ce dont mon

336. Claude Arnaud, «La voix de son maître», *Cinématographe*, N° 86 («Sacha Guitry»), février 1983, p. 10.

337. N. Burch, «Du muet, le parlant...», *art. cit.*, p. 60. Burch mentionne «les avant-gardes françaises» comme «Godard, Hanoun, Resnais, Robbe-Grillet, Pollet».

338. Fr. Truffaut, «Préface», *art. cit.*

film se rapproche le plus, c'est *Le Roman d'un tricheur* de Sacha Guitry.»<sup>339</sup> Or, à l'occasion d'un des moments importants de la réception du premier long métrage de Resnais, celui d'une table ronde réalisée par les principaux (futurs) représentants de la Nouvelle Vague dont les propos furent transcrits dans *Les Cahiers du cinéma*, Godard évoque Guitry, à l'étonnement de ses interlocuteurs dont plusieurs s'avèrent passablement déconcertés face à l'œuvre novatrice de Resnais :

Pierre Kast : « Cette alliance totale du film et de son scénario est si évidente que les ennemis du film ont aussitôt vu que c'est précisément là qu'il fallait attaquer : oui, le film est beau, mais il y a ce texte si littéraire, si peu cinématographique, etc. En fait, je ne vois pas du tout comment il est seulement imaginable de les séparer. »

Godard : « Tout ça ferait très plaisir à Sacha Guitry. »

Doniol-Valcroze : « Personne ne voit le rapport. »

Godard : « Si, le texte, le fameux problème du texte et de l'image. [...] Puisque Truffaut n'est pas là, je suis très content d'ouvrir à sa place une parenthèse pour dire qu'*Hiroshima* donne tort à tous ceux qui n'ont pas été voir la rétrospective Guitry à la Cinémathèque. »<sup>340</sup>

Si le film de Guitry et celui de Resnais diffèrent presque totalement du point de vue de la nature du texte, de l'expression orale, du ton et du mode d'instauration des relations verbo-iconiques (cf. chapitre VII), ils partagent une même propension à autonomiser la parole par rapport à l'image, de sorte que la voix-over apparaît, ici comme là, dans sa fonction de médiation. Alain Resnais réitérera d'ailleurs en 1980 la référence à Guitry dans un entretien avec Jacques Siclier réalisé à la sortie de *Mon oncle d'Amérique*<sup>341</sup>. De la part de ce cinéaste qui a constamment exploré les possibilités d'interactions entre la voix et l'image, la mention du film de Guitry prend une importance particulière. L'influence du *Roman d'un tricheur* me paraît être l'indice de la place de premier ordre que ce film occupe dans « l'histoire de la voix au cinéma ». Il se fait le

339. Cité par Robert Benayoun, *Alain Resnais, arpenteur de l'imaginaire*, Paris : Stock Cinéma, 1980, p. 81.

340. « Hiroshima, notre amour » (table ronde), *Cahiers du cinéma*, N° 97, juillet 1959, p. 6.

341. Jacques Siclier, « Conversation avec Alain Resnais. « Je ne fais pas un film consciemment », *Le Monde*, 22 mai 1980, p. 21.

représentant privilégié d'une phase intermédiaire, située entre le boniment des premières années du cinématographe et les pratiques d'expérimentation sonore de l'interrègne du parlant, dont certains cinéastes se réclameront un quart de siècle plus tard. Dans un entretien récent, Resnais confie d'ailleurs qu'il a été « influencé par la fausse conférence de Guitry au Théâtre de la Madeleine, intitulée *Ceux de chez nous*, où il faisait mine de lire une conférence »<sup>342</sup>. C'est probablement en raison du caractère oral de sa production cinématographique que Guitry a été le plus marquant.

### III.3 LE THÉÂTRE DE L'EXHIBITION VOCALE

Le *Roman d'un tricheur* est à compter parmi les rares films de Guitry à n'avoir pas été tiré d'une pièce de théâtre. Ce texte a par contre fait récemment l'objet d'une mise en scène théâtrale de (et avec) Francis Huster, représentée à partir de mars 2005 au Théâtre des Mathurins à Paris sous le titre *Les Mémoires d'un tricheur*. Le passé n'y étant pas « audiovisualisé », mais intégralement relaté par un narrateur qui s'adresse à un narrataire prétexte, cette version scénique se situe volontairement à l'opposé de l'activité mimétique traditionnellement associée à la représentation théâtrale, qui postule la monstration des actions effectuées par les personnages. Huster opte significativement pour la mise en scène du conteur lui-même, donnant ainsi chair à une présence qui s'affirme principalement dans le film par le biais de la voix-over: « À la voix extérieure du film métallique et sublime succédera au théâtre la voix sobre intérieure et fragile du tricheur incarné. »<sup>343</sup> La représentation s'achève par un enregistrement radiophonique de Guitry, ultime hommage à la « voix du maître ».

L'adaptation théâtrale tardive de ce texte ne signifie pas que tout indice de « théâtralité » soit absent du travail cinématographique de Guitry dans les années 30. On pourrait dire en effet que l'importance qu'il accorde à la voix est sans conteste liée à la passion qu'il nourrit envers le théâtre, attisée par son admiration pour son père comédien. Ainsi François Truffaut soulignait-il à juste titre combien les préoccupations de Pagnol et Guitry n'avaient pas seulement trait à la parole, mais aussi aux phénomènes de vocalité proprement dits :

342. Suzanne Liandrat-Guigues et Jean-Louis Leurat, *Alain Resnais. Liaisons secrètes, accords vagabonds*, Paris: Éditions des Cahiers du cinéma, 2006, p. 179.

343. Francis Huster, *L'Avant-scène théâtre*, N° 1181, 2005, p. 6.

« En tant qu'hommes de théâtre, [ils] étaient entraînés à privilégier la parole donc la voix, donc le rythme, donc les silences, donc la bande sonore. »<sup>344</sup> Alors que Pagnol s'intéresse prioritairement aux traits dialectaux susceptibles de camper un personnage en voix synchrone dans un environnement défini par son « coloris local », Guitry utilise une voix unique, la sienne, durant l'intégralité de son *Roman d'un tricheur* (à l'exception de la voix synchrone de la chanteuse du bar à soldats d'Angoulême, interprétée par Fréhel).

Guitry, marqué par son expérience théâtrale, compte sur sa voix pour instaurer une relation de proximité avec son public. Il ne faut donc pas concevoir la « théâtralité » de ses films comme une composante « non (voire anti-)cinématographique » (sous peine d'adhérer à un discours sur la spécificité du cinéma qui s'accorde mal à une problématique intermédiaire), mais plutôt comme un moyen de placer un médium fondamentalement associé à l'écrit sous le signe de l'oralité. La récurrence des questions liées à la représentation théâtrale chez les auteurs qui ont pensé les phénomènes d'oralité tient foncièrement au constat établi par Erving Goffman selon lequel, « profondément incorporées à la nature de la parole, on retrouve les nécessités fondamentales de la théâtralité »<sup>345</sup>. C'est précisément en tant qu'artiste de la scène que Guitry fait en quelque sorte renaître la pratique bonimentorielle, disparue en Occident et qui touche à sa fin à cette époque au Japon, là où, selon Lacasse, le film constitue encore une « attraction externe encadrée dans le spectacle théâtral »<sup>346</sup>. Sacha Guitry entreprend une perpétuation identique de certains aspects de la tradition théâtrale, mais les combine avec une acceptation de la voix enregistrée. Son cinéma relève assurément du « théâtre », non pas dans son opposition au cinéma, mais dans le sens large que Paul Zumthor donne à ce terme<sup>347</sup>. Il est vrai que le Guitry-locuteur du *Roman d'un tricheur* ne vise jamais à simuler de façon réaliste une interaction verbale : comme le note Anne-Sophie Nedelec, le cinéaste, qu'elle qualifie de « conférencier qui donnerait une leçon sur la manière dont on joue », « utilise sans cesse une voix de théâtre, impression renforcée par l'usage du ver blanc » et « use des procédés phatiques d'une manière quelque peu artificielle »<sup>348</sup>.

344. F. Truffaut « Préface », *art. cit.*, p. 17.

345. E. Goffman, *op. cit.*, p. 10.

346. G. Lacasse, « Du cinéma oral... », *art. cit.*, p. 55.

347. Voir P. Zumthor, *Introduction...*, *op. cit.*, pp. 53 et 56.

348. Anne-Sophie Nedelec, « L'auteur en voix off », in Francis Vanoye (dir.), *Cinéma et littérature*, Paris : Université de Paris, 1999, p. 45.

Cependant, la théâtralité ostentatoire de la diction favorise paradoxalement chez cet auteur-acteur l'adhésion spectatorielle, l'auditoire reconnaissant la voix tantôt traînante, tantôt emportée de Sacha Guitry.

Lorsque, dans les années 20, Sacha Guitry prônait le théâtre au détriment du cinématographe, c'était notamment pour s'opposer aux cinéastes qui considéraient que « le charme de la voix est nul »<sup>349</sup>. La relation particulière qu'il entretient avec le public, en tant qu'auteur dramatique mais surtout en tant qu'acteur de théâtre, constitue pour Guitry l'intérêt premier de tout spectacle. Dans *Le Roman d'un tricheur*, il tente de créer ce même type d'intimité – simulée par rapport à sa version théâtrale – qu'il renforce en donnant l'impression que sa voix parle « par-dessus » les images, présentifiant le monde du film grâce à sa dimension vivante. Guitry entend imiter les spectacles *live* en recréant la *communion* qu'ils instaurent avec le public – cette expression que j'ai utilisée à propos de la médiation bonimentorielle (voir II.1) apparaît d'ailleurs sous sa plume: « En vérité, un comédien ne peut pas être admirable sans vous – il a *besoin* de vous, il a besoin de se sentir en communion avec vous. »<sup>350</sup> Cette préséance accordée à la relation avec le public explique pourquoi Guitry tient à user de l'ironie verbale au cinéma, un ton qui rapproche le locuteur du spectateur et permet de jouer la parole contre l'image, le présent de l'énonciation contre le passé du « ça a été » ciné-photographique. En effet, c'est en termes de temporalité que Guitry envisage l'attitude ironique:

« Vous devez penser, comme moi, que l'ironie est une des manifestations les plus précieuses de l'esprit. Or, ne vous êtes-vous pas aperçu que l'ironie des événements n'est évidente qu'avec un peu de recul? À la minute même où ils se produisent, ils n'ont pas ce caractère. Leur actualité les en prive. »<sup>351</sup>

On comprend que, pour Guitry, le cinéma est un moyen d'expression particulièrement propice à l'usage de l'ironie, puisque les actions qu'il représente ne sont jamais « actuelles » pour le

349. Sacha Guitry, « Une phrase étonnante », *Candide*, 18 février 1926, repris in André Bernard et Claude Gautéur (dir.), *Sacha Guitry. Le Cinéma et moi*, Paris: Ramsay, 1977, p. 47.

350. S. Guitry, « Pour le théâtre et contre le cinéma » [1933], repris in *Sacha Guitry. Le Cinéma et moi, op. cit.*, p. 74.

351. S. Guitry, « L'ironie au cinéma », *Paris-Soir*, 20 septembre 1936, repris in Sacha Guitry. *Le Cinéma et moi, op. cit.*, p. 82.

public du film. Dans *Le Roman d'un tricheur*, le rejet dans le passé des événements contés a pour corollaire l'exacerbation de la situation d'énonciation (la terrasse du café, avec Guitry et Marguerite Moreno), illusoirement affichée comme *présente*. La voix-over émane du *hic et nunc* pour commenter un passé révolu. Ainsi le narrateur déclare-t-il, à l'instant où l'image nous fait découvrir le hall de l'hôtel dans lequel son personnage travailla jeune homme : « La mode a bien changé depuis ce temps lointain. Toutes ces belles dames et tous ces beaux messieurs nous semblent un peu ridicules *aujourd'hui*, comme dans quarante ans on ne manquera pas de *nous* trouver un peu ridicules. » Non seulement le commentaire prend pour objet – en le pointant grâce à des déictiques spatiaux (« ces belles dames ») – un élément de l'image que le spectateur perçoit au même moment (l'habillement des figurants, élément d'une reconstitution qui n'en apparaît que plus « réelle » en étant présentée dans sa valeur documentaire), mais il inclut le spectateur (par l'intermédiaire de la narrataire diégétique) dans un pronom personnel pluriel (« nous », c'est-à-dire « vous et moi ») et l'ancre dans un présent (celui de la projection) grâce à l'évocation successive d'un passé et d'un futur. La voix-over donne le sentiment de provenir du même espace-temps que celui dans lequel évolue le spectateur. La narration rétrospective, qui souligne constamment la supériorité de l'énonciateur verbal sur le personnage visualisé, alterne avec des moments de participation directe aux événements, par exemple lorsque le narrateur se met dans une position qui est celle de son double de l'écran, s'exclamant au présent : « Dans quel pétrin me suis-je fourré ! » La superposition de ces deux modes entretient l'ambiguïté attachée à la fonction d'une instance qui intervient comme un intermédiaire entre les pôles du dispositif.

Dans la conférence qu'il donne à plusieurs reprises entre 1932 et 1933 sous le titre significatif de « Pour le théâtre et contre le cinéma », Guitry, adoptant un mode de communication oral qui se distingue en lui-même de la fixation de la performance des acteurs de cinéma, fait une « déclaration d'amour » au *public* de ses pièces qu'il différencie de cette *foule* qui « suit les matchs de football et va au cinéma »<sup>352</sup>. Ce *distinguo* entre public et foule témoigne, au-delà d'une conception élitiste de l'art, d'une volonté de marquer

352. Cité in *Sacha Guitry. Le Cinéma et moi, op. cit.*, p. 57.

une différence entre le cinéma et les spectacles vivants. Cette distinction vaut notamment pour la voix, d'une part fixée et réitérable à l'identique, d'autre part proférée *live* et soumise aux aléas du direct. Guitry énonce à propos de l'acteur de cinéma une maxime qui aurait déterminé sa démarche pour *Le Roman d'un tricheur*: «L'acteur que vous voyez sur l'écran ne joue pas, *il a joué.*»<sup>353</sup> Le cinéaste tient donc à ce que l'acteur soit, au sens de Zumthor, *l'auteur d'une performance*, comme il l'est sur scène; il souligne l'importance de la dimension aléatoire du spectacle.

On est donc en droit de penser que, après son revirement en faveur du cinéma, Guitry ait voulu retrouver la parole vive du théâtre dans cet autre moyen d'expression auquel il fut réfractaire en raison de la nature mécanique de son enregistrement. Pour ce faire, il «incarne» (en alternant incarnation et désincarnation) dans *Le Roman d'un tricheur* un narrateur extrinsèque aux images – comme le fut effectivement le bonimenteur du «cinéma» parlé – qui semble s'approprier le matériau filmé après coup, réinjectant du présent (celui de la narration, visualisé dans une situation-cadre dont les quelques occurrences ponctuent le film) dans des actions qui appartiennent à son passé de personnage. En mai 1936, Guitry s'exprime à propos de la perspective adoptée pour son nouveau film:

«J'ai cherché, et c'est parce que je prétends que, sur l'écran, l'acteur a joué que l'idée m'est venue d'écrire et de réaliser ce film nouveau qui va paraître. Ces images mouvantes, innombrables et diverses, seront les illustrations d'un livre, d'un roman, d'une vie racontée – racontée au passé. Le personnage qui sera sur l'écran ne dira pas: «Je suis malheureux aujourd'hui» – non, il ne dira rien. Il aura l'air malheureux, et la voix de celui qui raconte dira: «J'étais malheureux ce jour-là.» C'est tout.»<sup>354</sup>

Cette explication motive le choix du qualificatif générique de «roman» présent dans le titre du film. Toutefois, ce terme est également associé à un récit écrit, et rend de la sorte peu compte de la prédominance de l'oralité dans *Le Roman d'un tricheur*. Il me paraît donc intéressant d'examiner les rapports que le texte *over* du film entretient avec la version écrite qui lui a préexisté.

353. S. Guitry, «Pour le théâtre...», *art. cit.*, p. 72.

354. S. Guitry, «*Le Roman d'un tricheur*», *L'Illustration*, 30 mai 1936, repris in *Sacha Guitry. Le Cinéma et moi*, *op. cit.*, p. 138.



### III.4 MÉMOIRES ET ROMAN

Guitry, en bon « tricheur », brouille les cartes : la version scripturale du récit dont le film est une adaptation s'intitule *Mémoires d'un tricheur*<sup>355</sup>, alors qu'il s'agit en fait d'un roman autobiographique (fictionnel). En outre, ce texte qui fut rédigé deux ans avant la réalisation du film<sup>356</sup> et accompagné d'illustrations dessinées par Guitry comprend en sous-titre la désignation « roman », quand bien même il s'apparente plutôt par sa longueur au genre de la « nouvelle ». L'ambiguïté des catégories génériques profite au centrage sur la figure de l'auteur, puisque les termes « tricheur » et « mémoires » semblent dans un premier temps renvoyer à un être ayant réellement existé. Le ton mondain du narrateur *over* du *Roman d'un tricheur* fait d'ailleurs indubitablement allusion, pour le spectateur contemporain de la réalisation du film ou familier de l'œuvre du cinéaste, à la vie de dandy que mena Guitry en tant que personnage public. Un journaliste du *Daily Mirror* aurait même confondu le « je » fictionnel et l'auteur réel en publiant, ainsi que le rapporte Guitry, un article intitulé « Révélations sur la famille de Sacha Guitry »<sup>357</sup>.

Il faut remarquer que la version écrite intitulée *Mémoires* se caractérise également par une forte présence de l'oralité, non pas bien sûr en termes de vocalité, mais grâce à la création d'un « style oralisé ». Sylvie Durrer nomme ainsi un « artefact littéraire qui entretient des rapports fantasmatiques avec la communication orale » dont il rend fictivement compte<sup>358</sup>. Dans les *Mémoires*, l'imitation de la spontanéité et de l'interactivité de l'oral ne se confine pas aux situations dialogiques (par ailleurs réduites à la portion congrue), mais s'applique principalement à la « voix » du narrateur elle-même. En effet, celui-ci use à plusieurs reprises de formes interrogatives (« Qu'est-ce que Monaco? », « Alors, les Monégasques, où sont-ils? », « Les pillera-t-il? », « Parlerai-je de mes maîtresses? ») ou impératives (« Ne perdez pas une seconde, dites :

355. Je me référerai à l'édition des Presses de la Cité (1991) pour les mentions de pages. Notons que la formule « Mémoires d'un tricheur » avait provisoirement été choisie pour le titre du film, ainsi que le mentionne un contrat de production du 8 novembre 1935 conclu avec Sandberg de la (future) société Cinéa (selon Laure Parchomenko, « Le fonds d'archive de la société Cinéas », 1895, N° 43, 2004, p. 80).

356. Texte radiodiffusé par Radio-Cité le 18 septembre 1936, repris in *Sacha Guitry. Le Cinéma et moi*, op. cit., p. 143.

357. S. Guitry, « *Le Roman d'un tricheur* », art. cit., p. 144.

358. Sylvie Durrer, *Le Dialogue romanesque*, Genève: Droz, 1994, p. 39.

ici»), recourt fréquemment à une structure paratactique qui transparaît notamment dans les tirets (pratique associée depuis le roman du XIX<sup>e</sup> siècle à l'expression du dialogue), utilise constamment le pronom personnel « nous » (« Entendons-nous », « donc, inclinons-nous ») et s'adonne sans cesse à des digressions. En outre, on remarque un souci du phatique qui se manifeste notamment sous la forme de la particule « oui » avec laquelle débudent plusieurs paragraphes (« Oui, j'étais vivant parce que j'avais volé », « oui, c'est une opérette! ») ainsi que le quatrième chapitre (« Oui, revenons à Caen »). La présence du narrateur dans son énoncé, manifeste par la récurrence des verbes de parole, souligne combien le texte relève du *discours* au sens de Benveniste. Par contre, il ne présente pas les traits ouvertement associés à l'oralité tels que les marques phoniques ou morphosyntaxiques (à l'exception, pour ce dernier type, des asyndètes introduites par l'usage des tirets): Guitry tient probablement trop au « beau langage » pour « altérer » la langue.

L'intérêt que manifeste constamment Guitry pour la fonction linguistique basée sur le « contact » avec le narrataire s'avère encore plus marqué dans la version filmique, puisqu'il peut y actualiser le style oralisé, par exemple en transposant les points de suspension en pauses ou en hésitations, ou en soulignant les tournures exclamatives par l'intonation. Son film exploite pleinement les spécificités de l'oralité que la nouvelle devait se contenter de suggérer.

### III.5 GUITRY, « CONTEUR » AU SENS DE WALTER BENJAMIN

Dans *Le Roman d'un tricheur* où l'acte même de raconter est mis en abyme par un récit-cadre, la voix-narration est également liée de façon indéfectible au présent. Lors du passage à l'enchaînement, les mots tracés par le personnage dans le cahier disparaissent pour ne réapparaître que subrepticement, l'écrit s'effaçant devant l'oral. L'ancrage diégétique dans l'écriture du journal soumet donc le visuel au verbe. Il en va d'ailleurs ainsi de la genèse du projet, puisque, comme on l'a vu, ce film connaît une « source » écrite. Le visionnage ne permet pas d'établir si, comme le prétend Truffaut et plusieurs commentateurs après lui<sup>359</sup>, une bande-son pré-établie servit de guide pour tous les gestes au moment du tournage. En effet, comme le narrateur fait porter de nombreuses

359. F. Truffaut, « Préface », *art. cit.*, p. 19; D. Desanti, *op. cit.*, p. 208.

remarques sur des actions profilmiques anodines et purement circonstanciées, il semble peu probable que toutes les composantes de la bande sonore aient précédé le tournage et le montage-images. Quoi qu'il en soit, ces moments ponctuels de référence simultanée à une action de courte durée et dépourvue de pertinence narrative nous donnent l'impression que le narrateur commente les images qu'il a devant les yeux, et qu'il en est donc également le (premier) spectateur, à l'instar du jeune Africain de *Moi, un noir* (Jean Rouch, 1958). La proximité instaurée avec le public du film présente par conséquent une parenté avec le dispositif du cinéma parlé. La précision du synchronisme est pour beaucoup dans la création de l'effet de présence du narrateur, car elle suppose un rapport d'immédiateté entre la profération et le défilement des images. Citons par exemple le procédé récurrent du décompte verbal progressif et simultané d'éléments diégétiques (les douze convives à table, les huit sous que l'enfant dérobe), la répétition du mot « cinq » lorsque le Russe indique ce chiffre de sa main ouverte, les diverses salutations échangées avec des personnages croisés dans la rue, et quantité de réactions « sur le vif » liées à une action en cours, par exemple lorsque la compagne du tricheur saisit un tableau suspendu au mur (« Voilà qu'elle déménage, maintenant ! »), lorsqu'elle est étendue dans la baignoire et lui tend sa jambe (« Donne-moi tes pieds d'abord, toi là ! ») ou lorsque Guitry jette un regard sur les manches de sa chemise (« Oui, elles iront encore demain »).

Tous les indices d'oralité relevés jusqu'ici confèrent à l'instance d'énonciation verbale du *Roman d'un tricheur* le statut qu'attribue Walter Benjamin au « conteur » dans un article qu'il a par ailleurs écrit l'année même de la sortie du film de Guitry<sup>360</sup>. À cette époque, l'auteur allemand dit constater un déclin, parmi les auteurs contemporains, d'individus véritablement capables de raconter une histoire. À de nombreux égards, le film de Guitry semble répondre à certaines exigences posées par l'essai de Benjamin qui, bien que consacré à l'écrivain russe Nicolas Leskov, propose plusieurs hypothèses théoriques qui ont une validité plus générale.

360. Walter Benjamin, « Le conteur » [1936], in *Œuvres III*, Paris : Gallimard, 2000. Bien que Benjamin ait lui-même traduit le titre original « Der Erzähler » par « Le narrateur », j'opterai également, comme les responsables de l'édition utilisée, pour le terme « conteur » afin de dissocier cette notion de l'usage dominant en narratologie littéraire et filmique.

Pour Benjamin, le conteur appartient aux pratiques traditionnelles de l'oralité, il est un pourvoyeur de « récit » – que Benjamin oppose au « roman » – c'est-à-dire qu'il relève du « domaine de la parole vivante »<sup>361</sup>. La fonction principale du conteur consiste à rendre verbalement transmissible une *expérience* que l'on trouve parfois dans le roman d'éducation. Or, d'une certaine manière, l'histoire narrée par Guitry s'inscrit dans ce genre: l'auteur y raconte bien les différentes étapes de la vie du personnage qu'il interprète. La description proposée par Benjamin s'applique tout à fait au récit à la première personne du *Roman d'un tricheur*:

« *Le conteur imprime sa marque au récit, comme le potier laisse sur la coupe d'argile l'empreinte de ses mains. Les conteurs ont toujours tendance à rapporter en premier les circonstances dans lesquelles ils ont entendu ce qu'ils s'approprient à raconter, quand ils ne le présentent pas simplement comme quelque chose qu'ils ont eux-mêmes vécu.* »<sup>362</sup>

On l'a dit, peu de cinéastes se soucient comme Sacha Guitry d'imposer leur présence (physique, vocale, nominale) dans leurs films. Dans *Le Roman d'un tricheur*, les retours à la situation de la terrasse du café permettent au conteur de se présenter et de présenter son récit. Ils autorisent également une dimension commentative que Benjamin attribue au conteur: tout récit « présente toujours, ouvertement ou tacitement, un aspect utilitaire »<sup>363</sup>, c'est-à-dire une capacité à délivrer un enseignement. Les digressions auxquelles s'adonne Guitry à propos des revers du destin ou de cette infamie qu'est la thésaurisation relèvent, en dépit de leur valeur ironique, de la « leçon de vie » transmise par le conteur benjaminien. C'est pourquoi son film ne ressortit pas, en dépit de son titre, au « roman » tel que le conçoit Benjamin, mais bien au « récit ». Lorsque, dans des films ultérieurs, Guitry abordera la grande histoire, il le fera également sur le mode de la narration orale décrite par Benjamin, c'est-à-dire en se faisant le *chroniqueur* des faits. Au début de *Si Paris nous était conté...* (titre qui renvoie d'entrée de jeu au rôle du *conteur*), nous découvrons après une entame en voix-over l'« auteur » assis à son bureau, lisant un livre d'histoire à de jeunes

361. *Ibid.*, p. 120.

362. *Ibid.*, p. 127 (je souligne).

363. *Ibid.*, p. 119.

gens. À la demande de l'un de ses auditeurs, il abandonne la source écrite qui devient selon lui ennuyeuse à force de précautions scientifiques pour proposer sa propre version des faits, en précisant :

« Ce sera le Paris, le Paris d'autrefois, mais regardé avec des yeux d'aujourd'hui – *et conté de mémoire*. [...] Oui, seulement je vous préviens, ma mémoire est fantasque... elle a ses préférences... elle est voyageuse. »<sup>364</sup>

La transmission verbale est dominée par le récit du chroniqueur qui, comme le note Benjamin, « ne vise pas à enchaîner rigoureusement les événements les uns aux autres », mais à en fournir une interprétation<sup>365</sup>. Guitry prend plaisir à mettre en scène la situation même de la communication d'un savoir qui, dans *Si Paris nous était conté*, s'effectue d'une génération à l'autre. Comme dans la plupart de ses films, Guitry corrèle fortement « narration » et « mémoire ». Benjamin rappelle combien la transmission orale des récits s'origine dans l'activité mémorielle du conteur : « Mnémosyne, celle qui se remémore, était pour les Grecs la muse de l'épopée. »<sup>366</sup> Dans *Le Roman d'un tricheur*, la destinée individuelle du héros croise d'ailleurs l'histoire collective sous la forme d'images documentaires ou d'archives insérées dans le film. La voix-over occupe alors une position similaire à celle qu'on lui octroie généralement dans des films non fictionnels ; tel est par exemple le cas de la séquence consacrée à la ville de Monaco, une sorte d'*À propos de Nice* (1930) auquel on aurait ajouté un commentaire verbal en actualisant l'ironie que connote le montage du film de Jean Vigo.

On le voit, la représentation de l'instance narrative du *Roman d'un tricheur* correspond en de nombreux points à la conception développée par Benjamin du représentant de la tradition orale. Ce rapprochement concerne même le ton du film, puisque Benjamin affirme à propos d'un passage humoristique de Leskov que l'« on voit percer ici la sympathie traditionnelle des conteurs pour les filous et les escrocs »<sup>367</sup>. L'apologie de la tricherie à laquelle Guitry consacre son film, le détachement qu'il affecte par rapport aux questions de légalité et de moralité soulignent cette insouciance du

364. S. Guitry, *Sacha Guitry. Cinéma, op. cit.*, p. 1076.

365. W. Benjamin, « Le conteur », *art. cit.*, p. 133.

366. *Ibid.*, p. 134.

367. *Ibid.*, p. 146.

conteur, tout entier livré aux plaisirs de la transmission orale. On retrouve ce détachement et cette insolence dans son dernier film, *Assassins et voleurs*, également construit sur un enchâssement narratif et une voix-over, à propos duquel François Truffaut avait souligné à sa sortie en 1957 «l'immoralité d'un scénario et d'un texte cynique glorifiant l'adultère, le vol, l'injustice et l'assassinat»<sup>368</sup>. *Le Roman d'un tricheur*, en dépit de la forme romanesque dont son titre semble faire un programme, est fait d'un régime vocal hybride qui, à certains égards, donne l'illusion d'entendre à nouveau la «voix-attraction» du cinéma parlé.

### III.6 LA REFLEXIVITE COMPLICE D'UN GÉNÉRIQUE PARLE

On sait que les premières images d'un film exercent une incidence considérable sur la lecture du spectateur dans la mesure où celui-ci se forge, à partir de ces informations liminaires, un certain horizon d'attente. Le générique, en tant qu'élément «péritextuel», participe nécessairement de ce réglage pragmatique, les informations qu'il convoie suscitant nombre d'inférences intertextuelles. Le générique du *Roman d'un tricheur* est à cet égard doublement particulier: il est à la fois *parlé*<sup>369</sup> et (partiellement) *diégétisé* (le monde du film se met déjà en place à l'image). De plus, les mentions écrites n'y sont pas simplement converties en phrases préférées comme dans *Le Mépris* (1963) de Jean-Luc Godard (autre cinéaste qui a fait de l'inventivité de ses génériques une sorte de «marque de fabrique»), mais supplantées par un usage du langage qui relève ouvertement de l'oralité. Quant à la diégétisation, elle ne fonctionne pas sur le mode d'une plongée immédiate dans l'univers du film, puisqu'on nous montre tous ceux qui ont participé à sa réalisation, techniciens compris. Ces deux aspects convergent dans l'usage de la voix-over qui interagit avec les individus filmés comme si elle provenait d'un espace proche. Traditionnellement, le degré de réflexivité d'un générique est élevé, mais Guitry souligne plus encore que dans la majorité des films l'origine du discours, qui est aussi celle du film: «Ce film, je l'ai conçu et

368. F. Truffaut, «Du cinéma pur», *Cahiers du cinéma*, N° 70, avril 1957, p. 46.

369. Guitry réalisera par la suite plusieurs autres génériques parlés. Dans quatre des cinq films qui succèdent au *Roman d'un tricheur* dans sa filmographie, le générique est présenté par le cinéaste en voix-over (*Faisons un rêve* et *Le Mot de Cambronne* en 1936, *Les Perles de la couronne* et *Désiré* en 1937).

je l'ai réalisé moi-même» est la première phrase *over*. Autant le « je » s'affirme verbalement (et même à trois reprises dans le même énoncé) ainsi que dans le grain si reconnaissable de la voix nasale du cinéaste, autant Guitry est le grand absent de cette présentation visuelle des différents protagonistes. Le second plan, succédant à celui du titre – qui est déjà une image iconique animée (quelqu'un retourne des cartes à jouer pourvues de lettres qui constituent progressivement le titre) – est révélateur de ce mouvement contradictoire d'exhibition et d'effacement qui se réalise pleinement dans la présence-absence de la voix-over. Il nous montre Guitry en amorce, filmé de dos, traçant sa signature sur une grande pancarte; lorsqu'il en a terminé, l'image défile dans le sens inverse et le mot disparaît. La même démarche sous-tend le recours à une voix sans corps: elle permet à l'auteur de s'exposer dans ce qui le caractérise (le grain de sa voix, son intonation, son humour, etc.) tout en le soustrayant, en tant que narrateur, à notre regard. Ainsi la voix-over mentionne-t-elle les principaux intervenants sans prononcer le nom de Guitry, ni en tant que réalisateur (ce dernier étant en quelque sorte compris dans la première phrase) ni en tant qu'interprète. Le plan de la signature encadre le générique, car il est suivi d'une seconde démarcation qui définit le « véritable » début du film. Il s'agit du moment où Guitry-personnage (de conteur), portant un verre à sa bouche, nous est d'abord montré comme une silhouette filmée à travers une vitre, puis en pleine lumière dans le plan suivant qui raccorde sur le mouvement de son bras. « L'auteur » naît à l'image grâce au montage. Il n'est pas étonnant qu'il surgisse de l'ombre, qu'il apparaisse d'abord comme une figure abstraite découpée dans l'obscurité, comme issue des spectacles pré-cinématographiques. Le personnage provient de ce non-lieu de la voix-over pour, au gré d'un renversement spatial qui a tout d'un truc de prestidigitateur, passer de l'ombre chinoise à l'image ciné-photographique. Le générique est délimité par cette disparition du nom de l'auteur au profit de l'apparition physique de l'acteur.

Parmi les termes du « contrat de lecture » que pose ce générique, le principal a trait à la voix: son omniprésence et son omnipotence se manifesteront en effet sur l'ensemble du film. En associant la voix de Guitry à la figure du réalisateur, ce générique rapproche le narrateur diégétique de la source supposément humaine du discours filmique. Même si Guitry affiche ici les moyens techniques

nécessités par le montage, c'est l'humanité de sa voix qui nous donne accès à leur visualisation. Cette voix, unique, exclusive, c'est aussi un ton. Avant même que le récit ne soit amorcé, la présentation du générique esquisse la façon dont la narration sera menée: une certaine désinvolture (la voix-over hèle des personnages qui manquent à l'appel), une mise à distance ludique de la fiction (le commentateur dit de deux actrices qu'elles «auront bien de la peine à nous faire croire qu'elles ne savaient pas qu'on les cinématographiait») – une stratégie qui renforce l'inscription du «désaveu de l'effet fiction» propre à tout générique – et, surtout, une interaction constante entre l'instance *over* et le monde diégétique. Par exemple, tel un bonimenteur, Guitry fait mine de s'adresser à un personnage des coulisses en s'écriant: «Serge! Où es-tu, Serge?» Lorsque Serge Grave apparaît, Guitry double de façon synchrone les quelques exclamations monosyllabiques que la bouche de l'acteur articule clairement («Quoi?», «Ah!», «Bon!»), ce qui est d'une certaine manière «légitime» pour cet acteur-là puisqu'il incarne à l'écran le double de Guitry (son personnage adolescent). Ce type de rapport entre la voix-over et l'action visualisée se poursuit durant tout le film, bien qu'ici le pouvoir de contrôle de l'instance verbale semble mis en cause par l'attitude des collaborateurs, qui refusent apparemment de s'y soumettre. Si le présentateur *over* feint d'en savoir moins que ses «personnages» (il prétend par exemple ignorer où se cache Pauline Carton), c'est pour créer un «effet de réel» qui convient à la fonction d'immersion des premières images d'un film.

En dépit de cette relative insoumission des actions visualisées, on ne peut pas dire comme Roger Odin à propos de *Partie de campagne* que, dans *Le Roman d'un tricheur*, «le générique s'oppose au film». En effet, Odin met l'accent sur la dichotomie, dominante dans la majorité des génériques de films, entre deux régimes concurrents: le *lire* et le *voir*<sup>370</sup>. La suppression des mentions écrites au profit d'une voix-over crée au contraire dans *Le Roman d'un tricheur* une continuité entre le générique et le corps du film qui contribue à masquer le changement de niveau énonciatif. En outre, la modalité perceptive de l'*entendre*, qu'Odin ne rapporte qu'aux sons diégétiques, induit une immédiateté beaucoup plus

370. Roger Odin, «L'entrée du spectateur dans la fiction», in Jacques Aumont et Jean-Louis Leutrat (dir.), *Théorie du film*, Paris: Albatros, 1980, pp. 203-205.



grande avec le public que le texte scriptural extradiégétique qui, malgré sa composante iconique, s'affiche ouvertement comme non solidaire des images (textes en surimpression). Ce rapport direct au public persiste d'ailleurs même lorsque Guitry utilise des mentions graphiques (pour le titre du film et le nom du producteur) puisque celles-ci sont totalement diégétisées, données à voir par un « personnage » de l'image (une main retourne les cartes, le producteur ferme la porte). En dépit du contenu réflexif des paroles qu'elle profère, la voix-over constitue un facteur d'immersion diégétique du spectateur : alors que le texte médiatise l'accès au monde du film, la vocalité confère un sentiment d'immédiateté, notamment parce que cette voix est de prime abord familière au spectateur.

On peut rapprocher cette ouverture en voix-over de la volonté, manifeste dans les films des premières années du cinéma parlant, de prouver immédiatement qu'ils sont « 100 % talking ». Toutefois, si le générique parlé du *Mystère de la chambre jaune* (Marcel L'Herbier, 1930) peut en partie s'expliquer par une intention, l'exhibition technologique n'a plus tellement de raison d'être en 1936, alors que le parlant s'est imposé en France. D'ailleurs, si l'on compare le film de Guitry à celui de L'Herbier, il apparaît que le geste d'autodésignation consistant à « montrer le dispositif »<sup>371</sup> est bien moins marqué dans *Le Roman d'un tricheur*, qui met en avant l'interaction de la voix-over et des individus filmés plus que l'appareillage. Certes, le film de Guitry montre, dans leur environnement de travail, les opérateurs, les décorateurs, la monteuse et son équipe qui « posent » pour la caméra, ainsi que « Paul Duvergé [qui] en imprimait les sons, isolé dans sa cabine, comme en plongée ». Ces plans-là sont par contre fort brefs en comparaison de ceux que Guitry consacre aux acteurs, à la mise en scène qu'ils effectuent devant l'objectif. Alors que le générique du *Roman d'un tricheur* reste fondamentalement humain, celui de L'Herbier montre les individus de l'équipe comme écrasés par les machines qui occupent la plus grande partie du champ et dissimulent les visages. En outre, si Guitry feint un ton de camaraderie (il tutoie par exemple certains participants), la voix-over du générique du *Mystère de la chambre jaune* conserve le sérieux d'un ton officiel et ne renvoie pas à la singularité d'un auteur, mais au caractère

371. Voir Christian Metz, *L'Énonciation impersonnelle ou le site du film*, Paris : Méridiens Klincksieck, 1991, pp. 85-88.

impersonnel d'une instance de production («cette production que les films Osso ont l'honneur de vous présenter»), le cinéaste étant désigné de la même façon que les autres collaborateurs («cette production a été dirigée par Monsieur Marcel L'Herbier»). Cette comparaison révèle la particularité de la complicité que Guitry instaure avec ses acteurs d'une part, avec son public d'autre part. La voix-over du *Roman d'un tricheur* endosse indubitablement, dans le contexte de la voix fixée, le rôle de médiation qui était dévolu au bonimenteur.

*La voix* : le singulier n'est pas une vaine formule dans ce film où, à quelques rares exceptions près, nous ne percevons qu'une seule et même voix, celle de l'auteur. Je ne pense pas qu'il faille seulement voir dans cet indéniabre facteur d'unicité une manifestation du narcissisme d'un cinéaste qui aime à souligner qu'il a lui-même *signé* le film, comme il le fait concrètement au générique en exhibant sa signature manuscrite. En effet, en ramenant tout à son ego qui est aussi celui de son personnage, non seulement il crédibilise l'autofiction du tricheur (le récit au «je» autorise les introspections et les commentaires subjectifs), mais il crée également une forte complicité avec le spectateur qui est amené à le (re)connaître. Cette insistance à inscrire la présence de sa personne (et de sa personnalité) dans le film n'est peut-être pas sans lien avec une certaine angoisse de Guitry face à la désincarnation qu'implique l'enregistrement de la voix, puis sa diffusion *over*. Par-delà la «fragmentation» d'un personnage incarné par plusieurs acteurs à différents âges et qui, une fois adulte, se déguise en quantité d'avatars, la voix-over assure l'unité du sujet en faisant (fictivement) converger tous les niveaux énonciatifs du film, puisqu'elle demeure identique, qu'il s'agisse du générique, de la situation-cadre ou du récit enchâssé.

Le fait que le cinéaste nous montre, dans le générique, le preneur de son dans sa cabine n'est toutefois pas sans infléchir notre perception de la voix-over, un procédé lui-même soumis aux conditions de l'enregistrement phonographique. Toutefois, les opérations du tournage étant décrites au passé, le film tend à faire oublier au spectateur que l'énonciation verbale *over* est également concernée par cette technique. C'est paradoxalement en exhibant la fixation des sons que Guitry l'occulte et en conjure la dimension mécanique. Il en va de même d'un étrange dialogue dans un autre film que le cinéaste réalise l'année précédente, *Bonne chance* (1935), où il s'entretient dans une voiture avec Jacqueline Delubac. Il s'agit d'un

plan unique filmé en travelling avant, dans lequel nous voyons le capot du véhicule au premier plan, alors que ses occupants demeurent hors-champ :

(Claude) – Regardez bien la route en ce moment. Quand je ne vais pas plus vite que ça en conduisant, cela ne vous donne pas l'impression d'être au cinéma ?

(Marie) – Si, si, c'est très juste.

(Claude) – N'est-ce pas ? Et savez-vous comment les gens de cinéma s'y prennent pour faire ça ?

(Marie) – Non.

(Claude) – Eh bien ! il paraît qu'ils mettent tout simplement leur appareil dans la voiture.

[...]

(Marie) – Mais... les paroles qu'on entend...

(Claude) – Eh bien ! on m'a dit, les paroles, qu'on les enregistre ensuite au studio.

(Marie) – C'est bien invraisemblable.

(Claude) – D'ailleurs je dois vous l'avouer franchement, je ne l'ai pas cru.

« Je ne l'ai pas cru » : l'improbable candeur de cette assertion révèle le fondement d'une telle interruption métanarrative. En effet, il s'agit avant tout d'une question de *croissance* (en la fiction du film et en la « réalité » de la voix). Sacha Guitry, pour qui la voix vive est une composante essentielle du jeu de l'acteur, affiche l'artifice cinématographique de l'enregistrement phonographique et de la déliaison fondamentale pour mieux les désavouer, ainsi que le fait explicitement et en toute naïveté Claude, le personnage qu'il incarne lui-même. Dans ces premières années où l'homme de théâtre œuvre également pour le cinéma, l'opposition de la voix vive et de la voix enregistrée constitue une préoccupation centrale.

Pour conclure ces quelques remarques sur le générique du *Roman d'un tricheur*, je dirai que Guitry y exacerbe la fonction instauratrice du Verbe en faisant intervenir sa voix d'*entrée de jeu* (la fiction qui traite de la fièvre du jeu est elle-même un jeu). La parole fait advenir le film comme elle nous permet d'embrayer sur le récit enchâssé. Il s'agit là d'un rôle qui incombe très souvent à la voix au cinéma, mais ce qui est singulier dans *Le Roman d'un tricheur*, c'est que cette démarche vise à assurer la prégnance de la

transmission orale par l'intermédiaire de la voix-over. La voix fixée tente dans ce film d'échapper à l'*écrit* en simulant une performance orale à travers les différents procédés que j'ai relevés dans ce chapitre. Bien que systématiquement *over*, cette voix se situe à mi-chemin entre l'attraction et l'intégration narrative, deux régimes auxquels je consacrerai désormais la suite de cet ouvrage.

## IV. LES FONDEMENTS THEORIQUES DE LA VOIX-ATTRACTION

### IV.1 L'INTERMEDIALITE DE L'INTERRÈGNE

Afin de poursuivre mes investigations comparatistes sur le «cinéma» parlé et l'interrègne du parlant, périodes qui me paraissent productives du point de vue de l'exploitation du potentiel de la vocalité, je vais examiner désormais plus avant la notion de «voix-attraction» que je propose d'ériger en dénominateur commun. Il faut néanmoins préciser que le parlé et le parlant sont foncièrement différents en tant que dispositifs, le second étant au premier ce que l'écrit est à l'oral. La voix-attraction étant indéfectiblement liée à l'espace scénique, son intégration au film parlant relève plus de la référence intertextuelle que d'une véritable assimilation. Par conséquent, lorsque le terme peut prêter à confusion, je propose de considérer que la voix endosse dans le cinéma parlant une fonction «pseudo-attractionnelle», comme on parle de «pseudo-interactivité» pour qualifier des dispositifs audiovisuels qui, tout en affichant un appel à la participation spectatorielle, ne peuvent, dans les faits, susciter une réponse de la part du public.

En dépit de l'irréductibilité de la performance vocale à tout enregistrement phonographique, le cinéma parlant demeure susceptible, comme il me semble l'avoir fait avant la standardisation qu'il a subie au milieu des années 30, de s'inspirer des modalités spectaculaires du cinéma parlé afin de les transposer partiellement sous la forme de voix captées. L'étude du *Roman d'un tricheur* menée au chapitre précédent montre qu'il est possible d'établir des parentés entre le boniment et une pratique plus tardive comme la voix-over. Dans un texte où il discute le concept d'«intermédialité», Rick Altman a suggéré un rapprochement similaire.

«À partir de 1910, le cinéma figure comme tel dans le grand livre des médias [...]. Son nom désigne sans équivoque une projection d'images en mouvement accompagnées d'un piano ou d'un orchestre. Paradoxalement, c'est en devenant sonore qu'il perdra le droit de s'appeler «cinéma». Devenues bruyantes, mais parlantes, les images de cinéma ne sont plus perçues de la même façon; on ne les reconnaît plus. Dès lors, pendant la deuxième moitié des années vingt, le cinéma se trouve de nouveau en situation d'intermédialité.»<sup>372</sup>

En postulant, à l'instar d'Altman, d'importantes analogies entre certains usages spectaculaires de la voix du cinéma des premiers temps et le régime vocal qui s'installe temporairement au sein de la production des premiers *talkies*, je n'entends pas présenter le cinéma de la période 1927-1935 comme le résultat d'une «régression», car cette perspective reconduirait un *topos* regrettable des histoires «classiques» du cinéma<sup>373</sup>. Je propose au contraire de dégager de cette comparaison quelques grands paradigmes théoriques utiles à une réflexion sur les régimes vocaux du cinéma. À cet égard, ma démarche est redevable des thèses avancées par Gaudreault et Marion sur la constitution de chaque médium à partir d'une pluralité de séries culturelles, mais aussi des recherches historiques entreprises par Édouard Arnoldy sur ces deux mêmes périodes<sup>374</sup>. Si une telle comparaison est possible, c'est parce qu'à la fin des années 20 une nouvelle donne technologique surgit (le son-sur-film synchrone) qui favorise les relations du cinéma avec d'autres moyens d'expression. Il s'agit d'un moment d'«intermédialité» au sens d'Altman qui, pour être circonscrit de façon satisfaisante, exige que nous quittions le seul champ du cinéma.

Précisons que Rick Altman prend soin de distinguer l'intermédialité du «multi-média», ce dernier concernant selon lui le couplage volontaire de différents médias déjà totalement constitués dans des champs autonomes. Selon sa définition, l'intermédialité, résultat d'une «crise de la médialité» progressivement résorbée dans l'autonomisation médiatique, constitue un phénomène

372. R. Altman, «Technologie et textualité de l'intermédialité», *art. cit.*, p. 12.

373. Voir par exemple Jean Mitry, *Histoire du cinéma*, Vol. 4 («Les années 1930»), Paris: Éditions Jean-Pierre Delarge, 1980, pp. 15, 22 et 27.

374. André Gaudreault et Philippe Marion, «Un média naît toujours deux fois», in *Sociétés & représentations*, *op. cit.*; E. Arnoldy, *Pour une histoire culturelle...*, *op. cit.*

nécessairement provisoire. D'un point de vue historique, le cas du «cinéma» parlé est donc radicalement différent de tentatives menées à d'autres périodes de l'histoire du cinéma pour associer parole vive et projections animées, par exemple dans des installations où l'artiste intervient au sein d'un dispositif machinique. Le partage proposé par Altman entre «spectacle multimédia» et «intermédialité» pointe en fait les risques et les limites de notre démarche comparatiste. Si l'application du concept de «dispositif» peut laisser envisager une certaine transhistoricité des phénomènes, c'est toutefois en laissant de côté l'un des pôles de celui-ci : le spectateur. Comme on l'a dit à propos d'*Une séance Méliès* (voir II.6), il est tout à fait chimérique, dans le cas du cinéma des premiers temps, d'espérer «reconstruire» théoriquement ce spectateur historique, alors que nous avons totalement assimilé aujourd'hui la double résorption de l'intermédialité provoquée par l'hégémonie du «Mode de Représentation Institutionnel»<sup>375</sup> et par la généralisation du cinéma parlant. André Gaudreault, confronté à ces difficultés lors de sa collaboration à l'écriture du boniment des *Beautés du Québec* (Jean Guy, 1995), discute les implications méthodologiques de cette question en exposant le problème ainsi :

«Impossible, en effet, de «sauvegarder» le spectateur des premiers temps... Impossible de reconstituer «l'émotion esthétique» des premiers temps sans reconstituer, *à la fois*, l'objet et le sujet de ce rapport entre des images et des sons, d'une part, et des cerceaux, de l'autre.»<sup>376</sup>

Dans une optique de «reconstitution», le rapport voco-iconique est indissociable des circonstances de la réception, déterminées par des facteurs sociologiques, psychologiques, historiques, etc. Toutefois, l'importance de ce pan pragmatique n'empêche pas qu'une théorisation des relations entre images et paroles puisse se concentrer sur certains aspects du dispositif. Il me semble en effet possible de traverser certaines modalités de cooccurrences de la voix et de l'image distantes dans le temps comme autant de dispositifs

375. Selon N. Burch, *La Lucarne...*, *op. cit.*

376. André Gaudreault, «Distance et historicité. Problèmes de méthode de la «reconstitution» historique», in André Gaudreault, Germain Lacasse et Isabelle Raynauld (dir.), *Le Cinéma en histoire*, Québec; Paris: Nota Bene, Méridiens Klincksieck, 1999, pp. 120-121.

distincts, tout en gardant à l'esprit que les témoignages d'époque, à ce jour rarement exhumés, sont loin de combler toutes les lacunes.

À l'instar du cinéma des années 1895-1908 environ, celui du début du parlant se cherche, endosse de multiples désignations et interagit étroitement avec d'autres pratiques spectaculaires. Cette intertextualité apparue au moment des tentatives d'adjonction de la parole fut souvent associée par les historiens traditionnels du cinéma au retour d'une « théâtralité » propre aux bandes du début du siècle. Georges Sadoul débute par exemple le chapitre qu'il consacre aux années 30 par l'affirmation suivante: « Dialogues, chansons, musique imposaient un recours au théâtre. Il entraîna plus d'une erreur, mais il fut aussi indispensable qu'au temps d'un Méliès. »<sup>377</sup> Si la composante évaluative nuit au propos de l'historien, celui-ci témoigne néanmoins, à travers la comparaison avec Méliès, d'une conscience de la résurgence d'une certaine hétérogénéité des pratiques. D'ailleurs, Sadoul évoque par la suite l'opérette et la revue de music-hall, ce qui laisse supposer qu'il élargit le terme « théâtre » à tous les types de représentations scéniques. Dans sa théorisation de l'intermédialité, Rick Altman remarque que les premières comédies musicales reproduisent exactement ce que l'on trouve à la même époque sur les planches de Broadway<sup>378</sup>. Il n'est pas fortuit que son exemple fasse intervenir la composante vocale, car celle-ci constitue selon moi un élément déterminant de la situation intermédiaire engendrée par l'apport des techniques de reproduction et de diffusion du son. La voix-attraction est à la fois un point commun entre diverses pratiques spectaculaires, et, au cinéma, le lieu de mutations conséquentes induites par la reproduction phonographique, la synchronisation des sons avec la bande-image, l'amplification sonore adaptée à un spectacle collectif, etc. Dans un article antérieur, Rick Altman avait d'ailleurs noté que les développements technologiques de la période de transition au parlant convoquèrent successivement des modèles sonores qui appartenaient à des modes de représentation différents et que l'idéal de l'intelligibilité aurait été inspiré, avant de se modeler majoritairement sur le théâtre, par le « Public Address », type de discours défini par un rapport direct instauré avec l'auditoire<sup>379</sup>.

377. Georges Sadoul, *Histoire du cinéma mondial*, Paris: Flammarion, 1949, p. 238.

378. R. Altman, « Technologie et textualité de l'intermédialité », *art. cit.*, p. 18.

379. R. Altman, « Technologie et représentation: l'espace sonore », in Jacques Aumont, André Gaudreault et Michel Marie (dir.), *L'Histoire du cinéma. Nouvelles approches*, Paris: Publications de la Sorbonne, 1989, pp. 126-129.



Sans envisager une véritable « filiation » entre l'utilisation de la voix au début du parlant et certains usages antérieurs ou ultérieurs – ce serait postuler là une continuité par trop « idéelle » – j'aimerais relever des phénomènes de résonance entre deux périodes qu'il est possible de rapprocher du point de vue de la place conférée au « cinéma des attractions ». Il va de soi que ces liens sont des constructions théoriques qui définissent autant d'axes de pertinence : il ne peut donc être nullement question ni d'exhaustivité, ni d'une « vérité » historique qui serait de l'ordre d'un enchaînement causal. C'est pourquoi je propose d'asseoir l'approche historique sur une réflexion préalable d'obédience linguistico-pragmatique que je présenterai de façon synthétique sous la forme d'une typologie des manifestations vocales.

#### IV.2 LES REGIMES VOCAUX DU CINEMA

En adéquation avec le cadre théorique adopté jusqu'ici et en vue d'examiner la fonction attribuée par un film à la voix dans son rapport à l'audiospectateur, j'envisagerai les diverses voix du cinéma à l'intérieur de quatre régimes vocaux : la *voix-attraction*, la *voix-action*, la *voix-explication* et la *voix-narration*, le premier et le dernier types constituant deux pôles opposés qui définissent l'axe principal de l'analyse. Ces catégories, suffisamment générales pour englober et faire se croiser la voix extérieure au texte filmique et les voix enregistrées (synchrone et *over*), ne s'excluent pas mutuellement, mais, dans un film ou un extrait donné, l'une d'elles tend généralement à s'instituer en régime dominant. L'opposition entre *voix-attraction* et *voix-narration* est sous-tendue par une prise en compte du phénomène d'institutionnalisation dont le procédé de la *voix-over* est l'une des manifestations majeures.

La projection bonimentée s'avère particulièrement propice à l'exercice d'une *voix-attraction* dans la mesure où elle est le lieu de la « confrontation exhibitionniste » qui caractérise pour Tom Gunning le « cinéma des attractions »<sup>380</sup>. Le locuteur s'adresse directement au public, faisant parfois primer la fonction phatique (ou conative) sur la fonction référentielle, qui ne s'instaure qu'avec la mise en place (facultative) d'une diégèse. Toutefois, comme on l'a vu dans

380. Tom Gunning, « The Cinema of Attractions. Early Film, Its Spectator and the Avant-Garde », *Wide Angle*, Vol. 8, N° 3/4, 1986, p. 66.

l'étude d'*Une séance Méliès*, le boniment peut également se mettre au service du récit «filmique» dont il peut contribuer à assurer l'intelligibilité, par exemple grâce au titrage des parties. Les caractéristiques générales de la performance vocale induisent cependant, dans le cas du boniment, un potentiel attractionnel supérieur à celui d'une projection synchronisée mécaniquement.

Entre la voix-narration et la voix-attraction – objets purement théoriques qu'il s'agit d'appréhender comme des instruments analytiques – s'échelonnent quantité de cas de figure qui correspondent à autant de gradients sur l'échelle de la valeur attractionnelle ou de l'intégration narrative. Cette conception bipolaire, dérivée de l'opposition entre l'oral et l'écrit, a l'avantage d'offrir suffisamment de latitude et de généralité pour nous permettre d'aborder des *régimes* vocaux, c'est-à-dire des modes dominants régissant la configuration de dispositifs audiovisuels qui font intervenir la voix humaine.

Si le couple attraction/narration circonscrit les deux tendances majeures, il ne recouvre pas l'entière des usages courants de la voix au cinéma, d'une part parce que la narration ne constitue qu'un mode de mise en texte parmi d'autres, d'autre part parce que le régime dominant de la voix synchrone, caractérisé par la forme dialogique, ne ressortit pas au pôle attractionnel. C'est pourquoi je propose deux termes supplémentaires: la *voix-explication*, située au pôle de la voix-narration avec laquelle elle ne partage toutefois pas, sur le plan textuel, le même type d'organisation discursive; la *voix-action*, qui s'applique à toutes les occurrences vocales synchrones internes à la diégèse filmique.

Si la voix-explication ne présente, par rapport à la voix-narration, aucune spécificité dans la façon dont elle fait interagir les trois pôles du dispositif – elle se situe également du côté de l'écrit, mais peut intervenir dans des pratiques vivantes (les «conférenciers» en usaient fréquemment pour satisfaire à des exigences pédagogiques) – la voix-action se singularise par rapport aux deux pôles de l'attraction et de la narration. Ce régime vocal permet de rendre compte des voix attribuées à un personnage qui, visualisé à l'écran (ou du moins s'exprimant à l'intérieur de l'univers du film), véhicule une parole dont le contenu interagit de façon immédiate avec les actions montrées. La parole provenant de la diégèse est alors asservie, à des degrés variables (le discours verbal pouvant être plus ou moins performatif), à l'action qu'elle provoque ou dont elle procède.

Cette parole *en acte* est une sorte d'équivalent verbal des actions entreprises par les personnages. À l'instar de ces dernières, les paroles font avancer le récit de l'intérieur, en naturalisant son développement sans créer de situation d'énonciation diégétique qui, explicitement, endosserait une fonction narrative. Très majoritairement présente dans les scènes dialoguées dont l'élaboration a constitué avec l'arrivée du parlant le facteur déterminant des conventions scénaristiques<sup>381</sup>, la voix-action correspond à la description que Roland Barthes donne du dialogue. Lorsque le sémiologue illustre la fonction de relais assumée par le message linguistique dans son rapport à l'image, il décrit ainsi le rôle des situations dialogiques au cinéma :

« Rare dans l'image fixe, cette parole-relais devient très importante au cinéma, où le dialogue n'a pas une fonction simple d'élucidation mais où elle fait véritablement avancer l'action en disposant, dans la suite des messages, des sens qui ne se trouvent pas dans l'image. »<sup>382</sup>

La voix-action permet l'exploitation de ce qui constitue, pour Paul Zumthor, une qualité intrinsèque et ancestrale de la voix : « Dans sa fonction première, antérieure aux influences de l'écriture, la voix ne décrit pas ; elle agit, abandonnant au geste le soin de désigner les circonstances. »<sup>383</sup> Ainsi, lorsqu'un personnage est « extrait » de cette action pour devenir un narrateur, nous nous déplaçons sur l'axe voix-attraction *vs.* voix-narration en direction de la seconde, avec des degrés variables en fonction de l'importance du détachement qu'affecte le contenu des paroles par rapport aux actions dans lesquelles s'ancre la situation d'énonciation. La partition entre la visualisation des actions et celle des situations d'élocution (monologue et dialogue) peut varier considérablement. Le passage de la voix-action à la voix-narration se traduit souvent par l'abandon de la situation de profération (voix synchrone) au profit d'une audiovisualisation du contenu narratif (voix-over), et cela surtout lorsque la durée de la réplique est conséquente. Cette transition obéit à certains codes qui touchent autant à la prise de vues (cadrage plus rapproché du locuteur,

381. Voir A. Schwarz, *Der Geschriebene Film...*, *op. cit.*, p. 27.

382. R. Barthes, « Rhétorique de l'image », *art. cit.*, p. 45.

383. P. Zumthor, *Introduction...*, *op. cit.*, p. 54.

diminution de la profondeur de champ, etc.) qu'au traitement dont bénéficie la piste-son (suppression progressive des bruits diégétiques, privilège accordé à l'audibilité de la voix du narrateur, etc.). Bien que fréquente, cette pratique qui consiste à faire disparaître le sujet parlant au profit de la visualisation du référent de ses paroles n'est pas indispensable à la représentation filmique d'un personnage-narrateur. Toutefois, plus le récit verbal se prolonge dans le temps, plus l'audiospectateur s'attendra à l'emploi de la voix-over, celle-ci relançant l'intérêt pour l'image en montrant autre chose que la situation d'énonciation diégétique.

Parfois, l'accent est néanmoins mis sur le locuteur lui-même, dont les expressions faciales importent autant (voire plus) que le contenu des paroles. Dans *Une sale histoire* (Jean Eustache, 1977), où deux narrateurs se succèdent pour nous faire un récit identique, le cinéaste s'intéresse aux caractéristiques mêmes de la performance des conteurs, l'une étant improvisée, l'autre jouée par un acteur professionnel. Cet exemple pourrait laisser croire que la perméabilité de la frontière entre voix-action et voix-narration apparaît exclusivement dans des films situés en marge de la production dominante (Godard, Straub et Huillet, etc.). Pour démontrer qu'il n'en est rien, on peut évoquer un film hollywoodien «classique» comme *Sur la piste des Mohawks* (*Drums along the Mohawks*, John Ford, 1939), qui opère un ancrage insistant de la voix-narration dans l'image du locuteur Gil Martin (Henry Fonda).

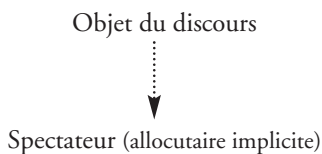
Dans le film de Ford, la bataille à laquelle le soldat Gil participe n'est pas montrée, le spectateur adoptant le point de vue de son épouse Lana (Claudette Colbert) qui reste à la maison. À une phase d'attente fait suite le retour de l'armée défaite par une nuit glaciale et pluvieuse. Les soldats, pour la plupart blessés, errent comme des fantômes et restent muets à la vue de Lana lorsque celle-ci s'enquiert auprès d'eux de la présence de son mari. Lorsqu'elle le retrouve dans une ornière, Gil se comporte différemment de ses camarades puisqu'il se lance dans un monologue ininterrompu, décrivant tout ce qu'il a vécu. Cette guerre que le film a tenue hors-champ n'existe pour le spectateur qu'à travers le témoignage verbal subjectif de ce soldat épuisé. Perdu dans son délire, tentant de surmonter son traumatisme en le verbalisant, Gil n'est mentalement plus «ancré» dans le présent de son énonciation. Il ne s'agit plus d'une voix-action dans la mesure où le locuteur ne tient nullement compte de ce qui l'environne (lorsque Lana s'éloigne pour chercher de l'eau, il poursuit

imperturbablement sa logorrhée), et n'exerce par ses paroles aucune influence sur l'action en cours, comme s'il parlait déjà d'outre-tombe. Allongé, le visage appuyé contre la paroi et bien visible, Fonda est filmé en plan d'ensemble après un léger travelling avant; il reste immobile et inactif alors que tout le monde s'affaire autour de lui pour soigner les blessés. Dans cette séquence, la prééminence totale de la parole synchrone s'accompagne d'un refus du découpage. Durant ce plan presque fixe qui dure plus de trois minutes (et se poursuit avec le même cadrage pendant une minute et vingt secondes après trois brefs inserts sur l'action parallèle d'amputation de la jambe du général), Gil fait au spectateur (plus qu'aux autres personnages qui l'écoutent à peine) le récit des actions qui ont été éliminées. Alors qu'une voix-over traditionnelle aurait été accompagnée des nombreux plans chargés de visualiser les actions décrites, cette voix *in* nous fait ressentir la présence du locuteur et la durée de son discours. Le poids de la temporalité de cette énonciation dramatisé la situation dans la mesure où nous nous demandons s'il ne s'agit pas là des derniers instants du soldat, s'il ne va pas s'éteindre avec ses paroles. La séquence ne s'achève d'ailleurs qu'au moment où il s'évanouit. Si la voix-narration convient aux instances *over*, elle connote également une certaine désincarnation lorsqu'elle est proférée *in*: le conflit entre la présence d'un corps et le retrait induit par la voix-narration exprime, dans cette scène du film de Ford, la situation critique d'un personnage à l'agonie. C'est en prenant le contre-pied de la représentation courante de la voix-narration que le cinéaste intensifie la séquence tout en la dédramatisant au niveau des actions représentées. L'éviction virtuelle de la corporéité du locuteur – la voix-narration étant généralement associée à une voix acousmatique – confère à cette scène la dimension hallucinatoire de la présence-absence.

Alors que la voix-narration tend à l'acousmatisation, voix-action et voix-attraction partagent une propension au synchronisme. Ces dernières se distinguent par contre en ce qui concerne la relation qu'elles établissent avec leur destinataire. En effet, la voix-action est adressée généralement à un autre personnage (l'aparté étant peu utilisé au cinéma où, justement, on préfère recourir à une voix-over pour rendre compte d'une intériorisation), tandis que la voix-attraction vise plus ou moins directement le public du film (ou, étant donné son statut pseudo-attractionnel dans le cinéma parlant, le public *dans* le film). La première renforce la clôture de l'univers diégétique, la seconde postule son éclatement.

L'ancrage de la voix-action dans l'espace-temps visualisé n'interdit pas qu'elle véhicule des informants narratifs<sup>384</sup>. Sylvie Durrer a souligné combien le dialogue romanesque diffère de la conversation courante (exception faite de sa dimension phatique) en montrant qu'il allie la logique de la conversation et la participation au récit: «Alors que la conversation est un phénomène d'essence sociale, le dialogue est une unité textuelle, dont les déterminations doivent être recherchées dans le contexte narratif.»<sup>385</sup> Cette soumission à l'organisation narrative concerne également la voix-action du cinéma. La visée de la voix-action est de ce fait analogue à celle de la voix-narration (ou de la voix-explication): toutes deux ont pour fonction de livrer au spectateur des informants sur le monde du film. Néanmoins, la voix-action s'inscrit dans une stratégie de naturalisation de cette transmission en s'adressant à un destinataire diégétique qui fait office de double du spectateur. Ce principe de délégation de la fonction communicative à une instance diégétique vaut pour le verbal comme pour l'image: il arrive en effet fréquemment qu'un film nous introduise dans une action par le biais d'un observateur dont nous sommes censés épouser le regard, de sorte que l'image, adressée à cet être de fiction, médiate notre relation à l'action. De nombreux informants étant convoyés par le canal verbal, il est nécessaire que les dialogues d'un film ne se limitent pas à représenter une relation «courante» entre deux individus, mais fournissent des données diégétiques et narratives placées dans les répliques à l'intention de ce tiers implicite qu'est l'audiospectateur. Par conséquent, la voix-action, en dépit de son autonomie diégétique fictive, est déterminée par trois instances, représentées dans le schéma suivant (la flèche pleine renvoie à une relation explicite, la flèche discontinue à une relation implicite):

Locuteur = personnage agissant  $\longleftrightarrow$  Allocutaire = personnage agissant



384. Voir sur cette notion A. Boillat, *La Fiction au cinéma*, Paris: L'Harmattan, 2001, pp. 125-126.

385. S. Durrer, *Le Dialogue romanesque*, op. cit., p. 66.

Cette double relation (diégétique et pragmatique) exige que les scénaristes (dont certains se spécialisent dans l'écriture des répliques) respectent deux critères de pertinence qui entrent parfois en contradiction: d'une part la cohérence diégétique – et donc la vraisemblance du dialogue – doit être maintenue tout en transmettant une information nouvelle au spectateur (un personnage qui, dans l'histoire du film, en connaît un autre depuis longtemps ne peut lui demander son prénom, à moins qu'une motivation narrative comme l'amnésie n'intervienne); d'autre part, on aura tendance à éviter de longues répliques qui rapportent une situation déjà connue du spectateur, même si l'allocutaire diégétique ne connaît pas cette information. Ces deux contraintes sont liées à une même nécessité de ne pas interrompre l'action, mais d'y fondre les informants narratifs utiles à la compréhension plus globale de l'organisation des différentes actions. Le cinéma classique hollywoodien, qui assure grâce à la redondance une lisibilité maximale du récit<sup>386</sup>, flirte bien souvent avec le premier cas tout en lui assurant une solide motivation. Il en va par exemple ainsi de la scène de la conversation entre les agents de la CIA dans *La Mort aux trousses* (*North by Northwest*, Alfred Hitchcock, 1959) qui, en introduisant pour la première fois dans le film un changement notable de régime dominant de focalisation (d'interne sur le héros, elle devient «spectatorielle» au sens de François Jost), est destinée à expliquer au spectateur que le personnage de Kaplan n'existe pas, qu'il n'était qu'un appât, une place vide que Tornhill est désormais venu occuper. Certaines motivations diégétiques (l'article de journal, la prise de parole du supérieur qui reformule l'ensemble des données, etc.) tentent de naturaliser ce transfert d'informants, mais la pertinence de plusieurs répliques au niveau diégétique reste faible. Ce type de rapport médiat au spectateur ne concerne pas seulement les films de fiction: le documentariste Frederick Wiseman opte par exemple systématiquement, dans ses films qui sont toujours dépourvus de voix-over, pour des situations où les gens dialoguent entre eux au lieu de s'adresser à la caméra sur le mode de l'entretien. Le cinéaste donne ainsi l'impression d'une clôture et d'une autonomie de l'univers diégétique, car ce sont les protagonistes agissants qui, sans renoncer à des activités quotidiennes qui conditionnent leur discours, nous livrent explicitement ou implicitement des informations sur l'institution à laquelle ils appartiennent.

386. Voir David Bordwell, *Narration in the Fiction Film*, London: Methuen, 1985, p. 57.

Inversement, il peut arriver qu'un dialogue n'ait aucune pertinence narrative. Par exemple, dans *L'Anglaise et le duc* (Eric Rohmer, 2002), le récit que la monarchiste fait à sa servante de la fuite qu'elle a entreprise de Paris à Meudon n'offre aucune information nouvelle : le spectateur vient de voir l'héroïne se cacher derrière un tronc d'arbre alors que passait un bûcheron, exactement comme elle le relate ensuite à sa confidente. Si Rohmer provoque cette redite, c'est pour mettre en évidence la « matérialité » verbale du texte qu'il adapte, *Ma Vie sous la Révolution* de Grace Elliott. Ce texte étant écrit au « je » (ainsi que l'indiquent les intertitres du film, composés de fragments des Mémoires), le cinéaste tient à le placer dans la bouche de la protagoniste interprétée par Lucy Russell, afin qu'il soit proféré comme il a été écrit et avec un accent anglais. Ici, c'est la source écrite qui légitime la propension de la voix-action à se faire voix-narration. Ce principe est appliqué de façon beaucoup plus radicale dans une œuvre antérieure du même cinéaste, *Perceval le Gallois* (1978), dans laquelle le rapport de domination que l'on observe traditionnellement entre voix-action et voix-narration est inversé, la première étant ici systématiquement neutralisée au profit de la seconde. En effet, il arrive fréquemment que les personnages de *Perceval* parlent d'eux-mêmes à la troisième personne, s'affichant comme les porteurs d'un texte qui leur préexiste. Les passages narratifs de Chrétien de Troyes n'ont pas été transformés en discours direct (ni d'ailleurs totalement en français moderne) en vue de leur intégration à l'action, mais conservent leur forme originale. Les répliques sont ainsi souvent autonomisées en raison de l'absence d'ancrage dans une situation dialoguée, « l'échange » apparaissant comme un usage de la parole qui est de pure convention. Par exemple, l'écuier reste totalement muet alors que Perceval lui adresse une longue tirade. En outre, la présence de la voix-narration est renforcée par les interventions d'un chœur (parfois visualisé, parfois hors-champ, mais dans tous les cas rapporté à des protagonistes privés de voix-action) qui assume la responsabilité énonciative (fictive) de certains moments de transition. En adéquation avec la démarche anti-illusionniste que le film affiche au niveau profilmique, la voix-narration est systématiquement associée à une voix synchrone, alors qu'elle supposerait une instance extérieure à l'action. Le statut énonciatif singulier des personnages leur confère le pouvoir d'influer sur la temporalité diégétique. Ainsi, Perceval résume ce



qu'il dit en même temps qu'il est censé le dire en usant du discours indirect («il explique que...»), déterminant ainsi un certain tempo narratif.

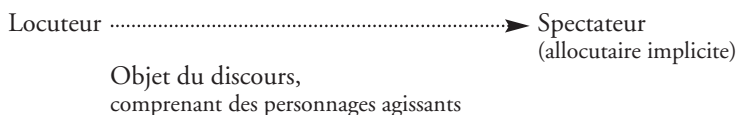
On le voit, la voix-action peut tendre à la voix-narration (ou à la voix-explication), et cela de façon inversement proportionnelle à l'importance de son ancrage dans la situation d'énonciation diégétique. Je tiens à rappeler que l'opposition entre la voix-action et la voix-narration ne repose pas sur l'absence (respectivement la présence) d'informants narratifs, mais sur la façon dont ils sont convoyés vers le public et insérés dans le flux visuel des actions. Les informants verbaux sont distribués parmi toutes les paroles du film, indépendamment de leurs rapports à ce qui est visualisé de la diégèse. Néanmoins, la voix-narration se démarque souvent à cet égard de la voix-action, tant du point de vue quantitatif (échappant à la contrainte de l'ancrage dans une source visible, elle se prête idéalement à la transmission d'une plus grande densité informative) que qualitatif, puisque le statut d'informant y est plus ostensible que dans les dialogues, où certaines phrases ne sont proferées que pour simuler une relation phatique. En outre, étant fictivement assimilé aux origines du discours filmique (voir chapitre VI), le locuteur *over* bénéficie d'une plus grande autorité qui vaut également pour ses dires.

Les situations d'aparté et certains commentaires dont le locuteur, situé à la frontière entre le diégétique et l'extradiégétique, se réfère à ce qui se passe simultanément autour de lui sans être solidaire de ces événements témoignent de la diversité des débrayages susceptibles de s'opérer à partir de la voix-action pour tendre soit vers la voix-attraction, soit vers la voix-narration. Je propose d'utiliser l'adjectif, emprunté à Christian Metz qui l'entend dans un sens quelque peu différent<sup>387</sup>, de «juxtadiégétique» pour décrire des situations où, à l'instar du bonimenteur qui s'exprimait en même temps que la projection, le personnage d'un film parle *pour* ou *de* quelqu'un d'autre qui se trouve dans le même espace-temps que lui, à l'instar, par exemple, du mari du personnage interprété par Liv Ullmann dans la séquence inaugurale de *Sonate d'automne* (Ingmar Bergman, 1978). La position particulière de ces personnages situés en retrait de l'histoire principale a l'intérêt de concilier la présence de la voix synchrone et la désincarnation de la voix-over.

387. Cf. Ch. Metz, *L'Énonciation impersonnelle*, op. cit., p. 54.

On trouve fréquemment de tels énonciateurs diégétiques chez René Clair, notamment dans *Les Grandes manœuvres* (1955), lorsque des figurants font des commentaires sur les protagonistes principaux qui se baladent à l'arrière-plan, ou dans la séquence de *Porte des Lilas* (1957) où le tenancier d'un bar lit à haute voix à ses clients un article de journal narrant les aventures d'un fuyard alors que, simultanément, les enfants miment dans la rue chacune des actions évoquées.

Sylvie Durrer note dans son étude des échanges verbaux qu'«introduire un récit dans un dialogue nuit dans la plupart des cas à sa fluidité et à son dynamisme, qui sont les deux justifications fondamentales de la présence du dialogue dans le roman»<sup>388</sup>. Il en va de même au cinéma, et de façon d'autant plus marquée que le dialogue s'appuie sur un jeu d'acteur dont on attend généralement qu'il soit «réaliste». L'intromission d'une «voix narrative» désincarnée qui s'opère à la faveur de la suppression de l'instance diégétique déléguée s'effectue au cinéma sous la forme de la voix-over, expression paroxystique de l'extraction du locuteur en dehors de son contexte d'énonciation visualisé. Si les caractéristiques vocales de la voix-action sont majoritairement déterminées par ce contexte (le débit sera par exemple l'indice de la précipitation ou de la crainte, etc.), la voix-narration se pare plutôt de la neutralité que lui confère sa position d'extériorité. La voix-narration – ou la voix-explication, qui ne diffère de cette dernière que sur le plan du mode d'organisation textuel – peut donc être représentée ainsi:



Christian Metz s'est montré particulièrement attentif à cette dimension communicationnelle de la voix-over qu'il qualifie significativement de «voix d'adresse hors l'image», précisant qu'«elle introduit le spectateur par un autre accès que ne fait le reste du film», «non point, en général, qu'elle <lui parle>, au sens ordinaire de ces mots, mais, encore un coup, elle parle pour lui»<sup>389</sup>. Cette

388. S. Durrer, *Le Dialogue romanesque, op. cit.*, p. 148.

389. Ch. Metz, *L'Énonciation impersonnelle, op. cit.*, p. 54.

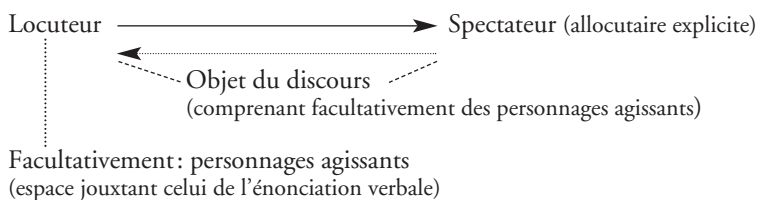
parole dégagée de toute contingence spatio-temporelle constitue une adresse qui, par convention, ne s'exhibe pas comme telle, mais que le spectateur reconnaît comme un canal privilégié de transmission de l'information narrative.

Bien que la voix-narration et la voix-attraction soient des régimes opposés, il se peut que des manifestations vocales empruntent simultanément certains traits à l'une et à l'autre. Dans le film *Perceval le Gallois* évoqué ci-dessus, la question du régime vocal se complique si l'on prend en compte la dimension chantée: le chœur nous conte les aventures du chevalier en musique, et certaines répliques sont proférées par les personnages sur des airs du XII<sup>e</sup> siècle joués sur des instruments d'époque. Il faudrait, dans de tels cas, distinguer parole-narration et voix-attraction, car si, habituellement, les caractéristiques vocales s'accordent à la fonction intégrative du texte, *Perceval* procède à une tentative de reconstitution d'une situation d'oralité où, à l'instar du boniment du cinéma des premiers temps, voix-attraction et voix-narration sont intimement imbriquées. Cette filiation avec certains «films» narratifs du début du siècle s'éclaire d'ailleurs à la fin du long métrage de Rohmer par une référence à un ancêtre commun, puisque le chevalier assiste dans une église à une représentation scénique muette d'une Passion du Christ, accompagnée de chants latins entonnés par des prêtres. Cette séquence finale d'environ huit minutes s'autonomise totalement du long récit développé jusque-là sur plus de deux heures, car *Perceval*, auquel le spectateur sait que l'ermite a conseillé de faire pénitence dans une église, n'apparaît pas parmi les spectateurs de l'épilogue. Or, comme le rapport entre les voix (celles des prêtres) et la scène (de la Passion) correspond à celui qui a régi l'ensemble du film, on pourrait dire que cette séquence constitue une mise en abyme de l'organisation énonciative de *Perceval*, dès lors rapprochée d'un mode de communication qui prévalait à l'époque représentée. Les voix de chacun des trois prêtres qui se dressent à l'avant-plan, filmés dans un espace que ni le cadrage ni le montage ne permettent de situer par rapport à l'espace scénique, sont prêtées respectivement à Jésus, à Ponce Pilate et au «narrateur» auquel incombent les «didascalies». Pour les moments non dialogués ou les paroles rapportées à un collectif (celui des prêtres qui condamnent Jésus), ces trois voix se joignent dans un chant choral qui ponctue les différentes phases du récit biblique. Ainsi le principe de dissociation entre acteur agissant et acteur parlant se voit inscrit dans une tradition de représentation

audiovisuelle qui remonte au Moyen Âge. Dans ce film, Rohmer use systématiquement d'effets de distanciation parents de ceux auxquels recourait Brecht pour ses pièces. J.-M. Valentin fait significativement remonter pour partie la conception brechtienne de la représentation à la tradition «du théâtre religieux, luthérien et catholique (jésuite), soucieux de proposer des modèles de vie, les *vitae sanctorum* prenant avec la Contre-Réforme le relais des Jeux de la Passion»<sup>390</sup>. La filiation avec ce type de représentations religieuses explicite les affinités entre le film de Rohmer et la pratique du boniment du cinéma des premiers temps. Dans un article sur les Passions filmées, Isabelle Raynauld mentionne d'ailleurs les Passions scéniques parmi les ancêtres de la projection bonimentée<sup>391</sup>.

C'est pourquoi *Perceval le Gallois* n'est pas sans présenter, au niveau du traitement de la voix, certaines analogies avec ce qu'ont pu être les projections bonimentées du cinématographe: voix-action et voix-narration sont émises par un même «acteur» au sein d'une performance globalement régie par l'esthétique des attractions. Comme on l'a vu, le bonimenteur avait également tout le loisir de passer d'une voix-narration à une voix-action en parlant à la place des personnages de l'écran. Tel était aussi le cas du *benshi* nippon, dont certains commentateurs définissent précisément le travail en évoquant ce point<sup>392</sup>.

En respectant les mêmes paramètres que ceux convoqués pour schématiser les deux premiers régimes, je propose de représenter ainsi la situation de communication induite par la *voix-attraction*:



La voix-attraction fait primer la relation directe au spectateur sous la forme de l'adresse, et peut suggérer la possibilité d'une interaction qui n'est effective que dans les spectacles vivants. «L'objet du discours» est secondaire lorsque domine ce régime vocal. À la

390. Bertolt Brecht, *Théâtre épique, théâtre dialectique*, Paris: L'Arche, 1999, p. 9.

391. I. Raynauld, «Les scénarios...», *art. cit.*, p. 137.

392. Voir Greenberg, *The Benshi...*, *op. cit.*, p. 8.

rigueur, le référent des paroles peut ne concerner que la relation au spectateur proprement dite. Alors que la fonction référentielle prédomine dans le cas de la voix-action et de la voix-narration, la voix-attraction fait passer au premier plan les fonctions phatique, émotive et conative liées à la communication de l'énoncé.

#### **IV.3 ATTRACTION/INTEGRATION (NARRATIVE): UNE CONCEPTION GRADUELLE**

La notion d'«attraction» que j'aborde ici sous l'angle de la voix est issue des hypothèses formulées par André Gaudreault et Tom Gunning à propos du «cinéma des premiers temps», champ privilégié des approches historico-théoriques. À l'occasion du colloque de Cerisy de 1985 consacré aux problèmes soulevés par l'histoire du cinéma, Gaudreault et Gunning introduisent dans leur intervention une opposition qui sous-tendra la plupart de leurs travaux ultérieurs sur la question<sup>393</sup>. Il s'agit de deux «systèmes» (ou «pratiques filmiques» au sens de Bordwell) apparus successivement durant la période 1895-1915: les «attractions monstratives» et «l'intégration narrative». L'attraction – terme que j'emploierai pour ma part au singulier (le caractère unitaire, individué de la voix s'accommodant mal de ce pluriel), sauf dans l'expression devenue courante de «cinéma des attractions» – se situe du côté du choc, d'une exhibition (voire d'un exhibitionnisme) qui est la source d'un plaisir perceptif pur. En termes d'organisation interne, l'attraction procède d'une juxtaposition d'éléments autonomes tels qu'on en trouve dans les spectacles de variétés; la narration vise quant à elle à créer une continuité, à intégrer l'ensemble des vues dans un tout où prime une logique de causalité. Dans une définition synthétique de la notion développée par Gaudreault et Gunning, Philippe Dubois énumère sous la forme de couples oppositionnels les principaux traits définitoires de l'attraction:

«Le «cinéma des attractions», c'est celui qui joue la monstration plutôt que la narration, la présentation plutôt que la représentation, la temporalité ponctuelle (*hic et nunc*) plutôt que l'organisation dans la durée, l'interpellation directe du spectateur

393. A. Gaudreault et T. Gunning, «Le cinéma des premiers temps: un défi à l'histoire du cinéma?», in *Histoire du cinéma, nouvelles approches, op. cit.*

plutôt que sa suspension et son conditionnement au suivi d'une diégèse, l'exhibition accrocheuse de ses propres moyens de figuration plutôt que l'effacement en vue d'une transparence de l'action racontée, l'excédent rhétorique plutôt que la norme classique, la recherche de l'événementialité visuelle «pure» par les formes les plus franches, etc.»<sup>394</sup>

Au vu de la généralité des aspects évoqués, on peut dire que le cadre théorique posé par la dichotomie attraction *vs.* narration – que l'on considérera plutôt comme une dialectique – peut s'appliquer non seulement à une analyse narratologique de certains films ultérieurs au cinéma des premiers temps, mais également, comme j'aimerais le montrer dans ce chapitre, à la question plus spécifique des rôles assignés à la voix au cinéma. Dans ce second cas, il est bien sûr nécessaire de transposer la description d'un spectacle visuel proposée par Dubois dans le domaine de «l'événementialité» vocale, la «monstration» se rapportant dès lors à la production d'un son, effectuée dans le *hic et nunc* d'une profération. Si la question du mode spectaculaire et de la temporalité s'avère également centrale pour aborder les messages linguistiques de la voix-attraction, traversés par un constant souci de présent(ific)ation, on verra que la dimension visuelle intervient également au niveau de la gestion de cet espace scindé entre le locuteur et son auditoire.

La notion d'«intégration» (narrative) proposée par Gaudreault et Gunning permet de créer un pont entre la période de l'attraction et celle de la narration. La prise en compte par ces historiens d'une étape intermédiaire – comprenant les films réalisés entre 1908 et 1915 environ – témoigne d'un souci de ne pas trop rigidifier la périodisation. La narration est en outre considérée comme une donnée soumise à un *processus* (d'intégration), non comme une forme canonique fixe. Toutefois, il faudrait disposer de la même latitude lorsque l'on distingue les pratiques attractionnelles et celles qui reposent majoritairement sur la narration en les concevant à l'intérieur d'un modèle graduel. C'est pourquoi il me paraît plus approprié de définir le *régime dominant* d'un film (ou d'un extrait de film) en fonction de la valeur attractionnelle ou du

394. Philippe Dubois, «La ligne générale (des machines à l'image)», in Frank Beau, Philippe Dubois et Gérard Leblanc (dir.), *Cinéma et dernières technologies*, Paris; Bruxelles: De Boeck, 1998, p. 84.

degré d'intégration narrative dont il fait preuve. Comme l'a argumenté Charles Musser – qui prône également un décloisonnement de la stricte opposition attraction *vs.* narration<sup>395</sup> – l'exclusion par Tom Gunning de toute narrativité durant la période qu'il identifie comme étant celle du «cinéma des attractions» ne résiste pas à l'examen de la narrativité interne des films de l'époque, et surtout pas à la prise en compte du potentiel créatif de leur contexte d'exploitation. Musser souligne précisément l'importance des conférenciers, dont le discours mettait en relation les différentes vues projetées, parfois en soumettant cette opération de cimentation à une logique narrative. Comme nous l'avons vu, la médiation bonimentorielles entre l'audiospectateur et la représentation pouvait tout à fait se baser sur des critères de narration semblables à ceux qui seront par la suite intégrés au discours filmique. L'historien note d'ailleurs que ces conférences «comportent de nombreuses similarités avec les documentaires à «voix off» des années 1950»<sup>396</sup> dans lesquels le verbal contribue à intégrer des images d'archives dans un ensemble discursif cohérent.

Le recours à une voix qui commente les films n'est toutefois pas le seul facteur de narrativisation susceptible d'apparaître au sein du «cinéma des attractions», bien que Gunning, dans une optique visant à rapprocher le cinéma des premiers temps de certaines pratiques d'avant-garde ultérieures, le définisse comme strictement non narratif. Musser mentionne avec raison l'exemple devenu canonique des Passions filmées déjà identifiées par Noël Burch et André Gaudreault comme ayant figuré parmi les premiers films à présenter un processus de linéarisation d'ordre narratif<sup>397</sup>. C'est d'ailleurs précisément à propos de Passions que Tom Gunning témoignera d'une conception graduelle qui n'apparaît généralement pas de façon aussi marquée dans sa discussion de l'esthétique des attractions: «Comme beaucoup de films des premiers temps, les Passions filmées déploient les énergies contradictoires de la période primitive en n'appartenant pleinement ni à un cinéma des attractions totalement dénarrativisé, ni à une trajectoire narrative qui serait en développement, mais présentant des aspects de ces deux tendances»<sup>398</sup>. Bien qu'il me semble

395. Charles Musser, «Pour une nouvelle approche du cinéma des premiers temps: le cinéma d'attractions et la narrativité», in *Les Vingt Premières Années du cinéma français*, *op. cit.*

396. *Ibid.*, p. 161.

397. N. Burch, *La Lucarne...*, *op. cit.*, chapitre 6 («Passions, poursuites: d'une certaine linéarisation»); A. Gaudreault, *Du littéraire...*, *op. cit.*, p. 30.

398. Tom Gunning, «Passion Play as Palimpsest: The Nature of the Text in the History of Early Cinema», in *Une invention du diable?...*, *op. cit.*, p. 107.

parfois enclin à minimiser l'importance de ces « énergies contradictoires », Gunning considère clairement, dans ce texte sur les Passions, que la conception bipolaire s'applique à l'ensemble de la période. Les Passions intéressent également Gaudreault parce que, bien que pluriponctuelles, elles convoquent exclusivement l'instance de monstration, c'est-à-dire qu'elles ne remplissent pas les conditions suffisantes pour qu'un film puisse, selon ses critères, être considéré comme le produit d'une narration. Gaudreault refuse ainsi d'y voir une activité narrative propre, et propose la notion de « narration extrinsèque » (dans ce cas, principalement véhiculée par un intertexte culturel) pour qualifier ce type de films où la juxtaposition de vues ne résulte pas de la structure filmique. Le présupposé du modèle de Gaudreault, induit par l'objectif de l'ouvrage qui réside dans une comparaison entre récits filmique, scénique et scriptural, est de l'ordre d'une « spécificité cinématographique » de la narration (alors qualifiée d'« intrinsèque ») : la narration n'apparaît pour lui qu'au niveau de la mise en chaîne<sup>399</sup>. Cette base théorique est tout à fait pertinente dans le cadre énonciatif qui est le sien, où il importe de désigner des instances (monstrateur *vs.* narrateur) et de discuter leurs fonctions respectives. Néanmoins, lorsque ce modèle est implicitement repris pour définir par défaut le cinéma des attractions, ce n'est plus seulement l'organisation filmique qui est en cause, mais aussi la mise en cadre et la nature du montré, qui concourent fréquemment à créer une situation voyeuriste. Or, à ce niveau-là, il semble difficile d'affirmer que tous les films réalisés avant 1906 sont dépourvus des traits communément associés au récit : l'événement profilmique induit un degré plus ou moins élevé de narrativité que le monstrateur peut accentuer en fonction du champ couvert par la caméra, du choix des actions filmées, du moment de l'enclenchement et du déclenchement de l'appareil de prise de vues, etc.

Ces différences légitiment une entreprise comme celle de Charles Musser visant à réhabiliter la propension au narratif de certaines vues. En fait, cette question de la « structure » narrative correspond précisément au pan de la narratologie que Gaudreault écarte sciemment dans son ouvrage *Du littéraire au filmique* afin de concentrer son étude sur les éléments propres au médium. Il ne faut pas oublier que les écrits sur l'esthétique des attractions sont

399. A. Gaudreault, *Du littéraire... , op. cit.*, p. 47.



informés par le contexte plus général de la narratologie filmique dans lequel André Gaudreault, partisan d'une définition très large de la narrativité – il n'identifie que deux conditions minimales (la succession et la transformation)<sup>400</sup> – occupait une position particulière. Selon sa conception, le seul mouvement provoqué par le défilement photogrammatique – à l'instar du bruissement des arbres à l'arrière-plan du *Repas de bébé* que les commentateurs de l'époque auraient si souvent mentionné – donne déjà lieu à du récit (mais non à de la « narration »). Si l'on compare cette bande, caractérisée par l'absence de moments marquants (on donne à manger à l'enfant du début à la fin), avec une autre vue Lumière d'avant 1900 intitulée *Premiers pas de bébé* (N° 67 du catalogue Lumière) qui est également filmée en un seul plan et présente un sujet parent (c'est le cas de le dire!), la nécessité de disposer d'une échelle de narrativité s'impose. En effet, dans *Premiers pas de bébé* qui montre (ou relate, raconte?) comment un bambin marche en titubant sur un trottoir à partir de l'arrière-plan pour venir, à l'issue d'une chute, effleurer de la main à la toute fin du métrage une poupée qui se trouve à l'avant-plan, on observe un nombre bien plus important de critères définitoires du récit: l'enfant est mû par une intentionnalité que le spectateur saisit immédiatement, et qui fait de lui un sujet au sein d'un schéma actanciel complet; le temps s'articule sur l'espace à parcourir, qui correspond, comme dans *L'Arroseur arrosé*, aux limites du cadre; enfin cet espace est vectorisé par une action. Une micro-situation de suspense s'installe quant à l'issue de cette tentative basique de préhension. Certes, on rétorquera que la vue exhibe par là même une maladresse infantile touchante et convoque le contexte affectif du « film de famille », fonctionnant ainsi plus comme « stimulus » émotionnel que comme récit, surtout lorsque cette bande est comprise dans un mélange hétérogène de vues qui ne présentent pas toutes ces propriétés narratives. Toutefois, le « monstateur » a fait incontestablement en sorte de disposer la poupée bord cadre, de choisir une distance que le bébé parcourt en un temps équivalent à la durée exacte du métrage disponible, enfin d'opter pour un sol qui offre, juste avant la réussite de l'entreprise, un dernier obstacle. Dans cette brève action profilmique, il n'y a certes pas trace d'un *narrateur* au sens de Gaudreault, mais on y trouve des

400. *Ibid.*, p. 40.

indices certains de *narrativité*. C'est en ce sens que je comprends le besoin de relativiser le caractère absolu de l'esthétique des attractions, pour l'inclure dans une conception graduelle de la narrativité. On peut par conséquent affirmer que, dans la formule «intégration narrative», le premier terme compte plus que le second. En effet, la narration est avant tout considérée comme un processus d'assimilation unifiante de parties hétérogènes en un tout qui sera associé, après la légitimation artistique du médium, à la notion d'«œuvre». Ce processus permet de distinguer à la fois la période pré-griffithienne du cinéma qui lui a succédé, et, dans notre cas, le «cinéma» parlé du cinéma parlant.

Si l'on met l'accent sur la composante «intégrative», l'examen du pôle de la narration n'exclut pas les types d'organisation textuelle différents du récit. En effet, certaines caractéristiques discutées par Gaudreault et Gunning dans le cadre de l'opposition attraction *vs.* narration ne sont pas exactement du même ordre: l'esthétique des attractions engage un certain rapport entre le film et son spectateur qui s'inscrit dans une tradition de pratiques spectaculaires, alors que la narration concerne de prime abord un certain mode de mise en texte qui s'est développé dans une tradition écrite, où il côtoie bon nombre d'autres modes (descriptif, explicatif, argumentatif, poétique, etc.). Afin de dégager une plate-forme commune aux deux notions, il faut donc mettre l'accent sur les implications pragmatiques du récit, qui résultent précisément de sa fonction «intégrative», garante de l'«absorption diégétique» du spectateur<sup>401</sup>. Or, dans une certaine mesure, tous les modes d'organisation participent du même pôle que la narration: la logique discursive, quelle qu'elle soit, se distingue de la discontinuité des attractions. En outre, la narration peut exister concurremment à d'autres types organisationnels et entretenir avec eux différents liens hiérarchiques. Ce principe de coexistence vaut plus généralement pour les deux types de pratiques (narration/attraction). Il est donc possible de nuancer la prétendue «transparence» du cinéma dit «classique» sans occulter l'éventuelle part attractionnelle dont certains films de ce vaste corpus font preuve à travers l'utilisation de la musique, l'intrusion de moments chantés, le recours à certaines conventions de genres ou à une composante spectaculaire marquée.

401. Voir T. Gunning, «The Cinema of Attractions...», *art. cit.*, p. 66.

Dans l'expression « attractions monstratives », deux termes sont apposés qui s'avèrent quasi synonymiques, comme en témoigne l'utilisation autonome qu'en feront Gaudreault et Gunning après l'article qu'ils écrivent de concert. Gunning s'en tiendra à la notion d'« attractions » (plus appropriée à la langue anglaise) dont la paternité lui est attribuée par son collègue<sup>402</sup>, alors que Gaudreault, dans une perspective relevant de la théorie de l'énonciation narrative, développe l'opposition entre monstration et narration.

En établissant un lien entre monstration et récit scénique, Gaudreault renoue sciemment avec la notion d'« attractions » telle que l'a conçue S. M. Eisenstein dans un texte sur les représentations théâtrales qu'il réalisa pour le Proletkoul't<sup>403</sup>, et que Jacques Aumont décrit comme une « attraction de music-hall, un moment fort du spectacle, relativement autonomisable, faisant appel à des structures de représentations qui ne sont pas celles de l'illusion dramatique »<sup>404</sup>. La possible autonomisation des parties (le multiple *des* « attractions » s'oppose à l'unicité de l'« intégration ») postule une certaine résistance envers la logique cohésive du récit, voire la suppression de cette dernière. En définissant l'attraction au théâtre comme « tout moment agressif [...] soumettant le spectateur à une action sensorielle ou psychologique »<sup>405</sup>, Eisenstein met l'accent sur l'immédiateté du rapport au public, et donc sur une dimension pragmatique qui, reprise par Gunning, me semble également concerner la question de la voix. Toutefois, une différence de taille subsiste entre ces écrits récents et les théories du Soviétique: alors qu'Eisenstein corrèle montage et attractions, Gaudreault et Gunning postulent l'inverse. Pour eux, l'attraction consiste en une simple juxtaposition de « vues », tandis qu'Eisenstein opère une distinction entre la performance profil-mique, qui n'est qu'une « attraction présentée » (comme on dit « théâtre filmé »), et une attraction qui résulterait véritablement d'un travail réalisé essentiellement au niveau du montage<sup>406</sup>.

Sur la base de la notion d'« attractions », Tom Gunning développe ce qu'il avait appelé, lors du colloque de Brighton, le « style

402. A. Gaudreault, *Du littéraire...*, *op. cit.*, p. 32, note 15.

403. S. M. Eisenstein, « Le montage des attractions » [1923], in *Au-delà des étoiles*, Paris: Éditions des Cahiers du cinéma, 1974.

404. Jacques Aumont, *Montage Eisenstein*, Paris: Albatros, p. 57.

405. S. M. Eisenstein, « Le montage des attractions », *art. cit.*, p. 118.

406. S. M. Eisenstein, « Le montage des attractions au cinéma » [1925], in *Au-delà des étoiles*, *op. cit.*, pp. 128-129.

non continu du cinéma des premiers temps», dont les caractéristiques essentielles sont l'interpellation du spectateur par l'acteur et la présence, au lieu de la causalité du style continu, de liens fort ténus entre des «numéros» d'anthologie. Gunning souligne l'influence de la structure et des sujets du vaudeville sur cette période du cinéma en précisant que les films eux-mêmes faisaient partie de tels programmes<sup>407</sup>. Comme il n'est pas encore question parmi les historiens du cinéma des «sons du muet» à l'époque de ses travaux, Gunning n'intègre pas à sa réflexion la dimension sonore du spectacle, se contentant de mentionner en passant le fait que de nombreux films du début du siècle étaient conçus pour accompagner visuellement une chanson dont ils héritaient du titre. Il mentionne par contre un genre filmique qui présente à ses yeux certaines analogies avec la «non-continuité» du cinéma des premiers temps: la comédie musicale<sup>408</sup>. Gunning réitérera cette référence quinze ans plus tard en s'expliquant ainsi:

«Les attractions constituent un mode visuel d'adresse au spectateur et existent dans d'autres périodes que celle du cinéma des premiers temps. Certains genres comme la pornographie, la comédie musicale ou les actualités filmées restent étroitement liés à ces procédés tout au long de l'histoire du cinéma.»<sup>409</sup>

Même s'il indique clairement qu'il s'agit pour lui d'un «mode visuel», Gunning évoque trois genres qui tendent à conférer une place importante à la voix (expression du plaisir physique, support d'une chanson ou voix-over des actualités). La question de la valeur attractionnelle de la voix chantée concerne d'ailleurs en propre la période traitée par Gunning, comme en témoigne, aux États-Unis, la vogue des chansons illustrées à l'ère des nickelodéons<sup>410</sup>. Selon un procédé issu de la lanterne magique, les exploitants projetaient des *song slides*, c'est-à-dire des plaques sur lesquelles figurait le texte de chansons qui étaient entonnées par le pianiste et reprises en chœur par le public, à l'instar du karaoké actuel mais sous une forme entièrement *live*. On retrouve une citation de cette pratique

407. T. Gunning, «Le style non-continu du cinéma des premiers temps (1900-1906)», *Les Cahiers de la Cinémathèque*, N° 29, 1979, p. 26.

408. *Ibid.*, p. 33.

409. T. Gunning, «Cinéma et modernité», *Cinémathèque*, N° 5, 1994, p. 131.

410. R. Altman, *Silent Film Sound, op. cit.*, chapitre 10.

qui simule sous une forme fixée l'oralité qui régnait sur l'exploitation cinématographique des premiers temps dans *Strawberry Blonde* (Raoul Walsh, 1941), réminiscence nostalgique de l'époque dont il est question dans l'histoire du film.

Dans leur introduction à l'ouvrage collectif issu du cinquième colloque de Domitor, *The Sounds of Early Cinema*, Richard Abel et Rick Altman soulignent combien l'étude du son dans les salles contribue à « compliquer [la] notion de cinéma des premiers temps comme « cinéma des attractions »<sup>411</sup>, notamment du point de vue de la périodisation, puisque cette dominante attractionnelle se perpétue selon eux jusque vers 1910, voire plus tard. Dans ce cas, la voix est une composante de l'exploitation des films (le « cinéma » parlé), et non un élément constitutif du texte filmique.

Le chant que l'on trouve autant dans les pratiques vivantes du début du XX<sup>e</sup> siècle que dans les comédies musicales ultérieures me paraît intrinsèquement plus attractionnel que la parole parlée, à moins que cette dernière ne soit mise au service de sketches reposant sur une pure virtuosité langagière et fortement basés sur une interaction avec le public. En effet, l'exhibition des qualités vocales, l'intensité sonore importante, la soumission moindre du verbal aux motivations diégétiques et l'appel au spectateur (qui peut fredonner la chanson, bouger en rythme, etc.) expliquent le renforcement du pôle attractionnel de la parole lorsqu'elle est chantée. C'est pourquoi, chaque fois que Tom Gunning propose d'élargir la période des attractions au-delà du cinéma des premiers temps sans pour autant exclure le cinéma hollywoodien, il mentionne l'exemple du *musical*<sup>412</sup>, qui perpétue par excellence le pôle du « cinéma des attractions » au sein du parlant en conjuguant musique, chant et danse. La composante vocale est d'ailleurs évoquée incidemment à propos du cinéma des premiers temps par Gaudreault et Gunning, qui parlent de la présence d'un « showman » hérité du vaudeville et précisent, en note (ce qui est révélateur d'une volonté d'écarter cette composante dont ils ont néanmoins conscience), que l'activité « monstrative » comprenait aussi bien des « chansons » que des « sketches »<sup>413</sup>.

411. R. Abel, R. Altman (dir.), *The Sounds of Early Cinema*, *op. cit.*, p. xiii.

412. T. Gunning, « The Cinema of Attractions... », *art. cit.*, p. 64; « Cinéma et modernité », *art. cit.*, p. 131.

413. A. Gaudreault et T. Gunning, « Le cinéma des premiers temps: un défi à l'histoire du cinéma? », in *Histoire du cinéma, nouvelles approches*, *op. cit.*, p. 63.

En postulant l'existence d'un régime que je propose d'appeler la «voix-attraction», je n'entends minorer en rien l'importance des particularités visuelles du cinéma des attractions: le spectacle des images animées du «muet» s'offrait avant tout au regard du public. Toutefois, en prenant en compte les recherches récentes sur les sons qui accompagnaient les projections de cette époque, il me semble également possible de considérer sous cet angle la place octroyée à la voix dans le dispositif du «cinéma» parlé, et de corréliser ces observations avec l'étude de certains aspects du cinéma parlant. Le postulat d'une perpétuation de la voix-attraction durant le parlant exige par contre quelques ajustements liés à la dimension technologique du spectacle cinématographique. Pour éviter que la bipolarité attractions/narration théorisée à propos d'une période historique bien particulière soit transférée sans autre forme de procès du cinéma des premiers temps vers le cinéma parlant, il convient en effet de souligner l'importance épistémologique de la rupture qui s'opère avec la découverte et la généralisation du principe d'inscription de la voix. Les possibilités d'une synchronisation mécanique du son avec l'image ravivèrent en effet les aspirations anthropocentriques qui animaient déjà les férus de «machines parlantes» que certains inventeurs s'employaient à concevoir, notamment à partir du XVIII<sup>e</sup> siècle, lorsque la fabrication d'êtres ou d'animaux artificiels connut un véritable engouement. Il s'agit là d'une série culturelle et technologique dans laquelle vinrent sans nul doute s'inscrire pour partie les spectacles qui proposaient le couplage du phonographe et du cinématographe. La restitution phonographique se distingue pourtant fondamentalement de ces voix qui, avant Edison, étaient certes *produites* mécaniquement, mais non *reproduites*. Désormais la machine se fait le véhicule de l'humain.

Cependant, la trace n'est pas une présence effective: les vibrations des masses d'air qui ont été retranscrites appartiennent à un autre temps, voire à un autre lieu. Alors que, dans le cinéma parlant, la voix est *intégrée* au pôle machinique du dispositif, le bonimenteur est susceptible d'opérer, comme on l'a vu, une médiation entre les différentes instances qui interviennent dans le dispositif audiovisuel. La présence physique d'un locuteur dans l'espace de la salle et l'hétérogénéité globale du spectacle confèrent à la voix vive un statut foncièrement plus attractionnel que ne peut l'être toute voix enregistrée et diffusée lors d'une projection dont toutes les



d'un sujet parlant). Le second niveau d'intégration concerne l'organisation interne (généralement narrative) du film.

L'absence de premier niveau d'intégration dans le cinéma parlé ne pose pas de problème quant à la prise en compte de la représentation proprement dite. En effet, les circonstances de la projection et l'économie textuelle du « film » interagissent constamment : l'agencement libre des vues durant les premières années du cinématographe découle en grande partie du fait que l'exploitant achetait individuellement les bandes ; inversement, l'augmentation de la durée du métrage n'est pas sans incidences sur le mode de présentation des films. Le terme « attraction » réfère d'ailleurs autant à ce que montre le film (et comment il le fait) qu'à des pratiques spectaculaires qui ne font intervenir aucune projection d'images. La voix-attraction du cinéma parlé s'inspire des conditions qui ont régi le « cinéma » parlé, mais doit se contenter d'être une *reproduction* de celles-ci au sein d'un dispositif audiovisuel fondamentalement intégrateur. Ainsi, même lorsque le bonimenteur ordonne son texte selon les lois du récit, les circonstances de la réception de son discours confèrent à ce dernier un statut plus attractif que n'importe quelle voix d'adresse du cinéma parlé, où l'oralité est nécessairement médiatisée par la technique de restitution sonore.

Si j'ai opté pour conférer une place cardinale à la notion d'« intégration » en reléguant le « narratif » à une qualification facultative, c'est parce qu'à mon sens le pôle traditionnellement considéré comme étant celui de la narration se définit avant tout par la fonction intégrative conférée à la voix, non par le type discursif auquel ressortit le texte qu'elle profère. En effet, le primat accordé dans les productions cinématographiques courantes à la mise en forme narrative ne doit pas faire oublier, d'une part, que le degré de narrativité est variable – entre la « relation d'actions » et la « mise en intrigue », en passant par la forme élémentaire du micro-récit (parfois mis en série), l'organisation narrative d'un film obéit à une définition plus ou moins restrictive du récit<sup>415</sup> – d'autre part qu'il existe certains modes alternatifs d'organisation comme le descriptif, l'argumentatif, etc.<sup>416</sup>. L'évaluation de la place octroyée au narratif dans un film dépend en outre d'éventuels rapports de subordination ou de

415. Voir Alain Boillat, *La Fiction au cinéma, op. cit.*, pp. 183-191.

416. Voir par exemple la typologie des « systèmes formels non-narratifs » proposée par David Bordwell et Kristin Thompson dans *L'Art du film. Une introduction*, Bruxelles ; Paris : De Boeck, 2000 [1<sup>re</sup> édition en anglais : 1979], pp. 159-203.



domination par rapport à d'autres formes. Dans une étude portant sur l'hétérogénéité compositionnelle des discours à travers certaines organisations séquentielles types, Jean-Michel Adam aborde le récit au même titre que d'autres structures textuelles (la description, l'argumentation, l'explication et le dialogue)<sup>417</sup>. Une telle approche taxinomique ne mène pas seulement à isoler les différents « prototypes » à des fins de définition, mais aussi et surtout à concevoir des structures séquentielles hétérogènes en prenant en compte, ainsi que le propose Jean-Michel Adam, deux principes souvent interdépendants : l'*insertion* de séquences hétérogènes et la *dominante séquentielle*<sup>418</sup>. Il est alors possible de considérer des cas où un mode d'organisation dominant à l'échelle de l'ensemble du film laisse momentanément la place à un autre mode qui lui est subordonné. Par exemple, le type argumentatif instauré par le texte *over* de *Roger and me* (Michael Moore, 1989) est suspendu lors du récit autobiographique initial qui permet de linéariser des extraits disparates, jusqu'à ce que l'individuel rejoigne le collectif à travers l'interrogation que suscite la fermeture des usines dans la ville natale du cinéaste. La prise en compte de ces variations de types discursifs nécessite donc de penser la hiérarchisation induite par des phénomènes d'emboîtement. Or la voix-over constitue un moyen privilégié pour introduire des enchâssements, qu'ils soient narratifs ou qu'ils relèvent d'autres types. Les « séquences textuelles » *over* contribuent à organiser l'agencement des plans en obéissant à certains schémas prototypiques intuitivement identifiés par le spectateur, d'autant plus que le support linguistique permet de repérer (plus nettement que l'image) des connecteurs ou des procédures fonctionnant comme indices d'un type discursif particulier. Dans la grande majorité des films distribués dans les salles, cette diversité des modes d'organisation est toutefois le plus souvent réduite au mode dominant du narratif.

Notons enfin qu'il n'est pas possible, lorsqu'on tente d'évaluer le degré d'intégration seconde de la voix, de faire l'impasse sur la représentation visuelle du locuteur. Comme l'observe Édouard Arnould, « une histoire du cinéma sonore et parlant croise, en plusieurs de ses ramifications, une histoire du regard et des images en mouvement »<sup>419</sup>. Paul Zumthor a bien montré combien la transmission

417. J.-M. Adam, *Les Textes : types et prototypes*, Paris : Nathan, 1992.

418. *Ibid.*, p. 31.

419. É. Arnould, *Pour une histoire culturelle du cinéma*, *op. cit.*, p. 18.

orale est indissociable des conditions de la performance, qui comportent également des paramètres visuels. Au vu de son ancrage dans un espace donné (à l'origine scénique, mais qui peut être reproduit dans l'espace diégétique d'un film), la voix-attraction est fortement liée à la présence d'un corps. Plus celui-ci – et notamment le visage – est mis en évidence (par le cadrage, l'éclairage, le décor, etc.), plus la voix qui en émane peut prétendre au statut d'attraction ; moins la représentation filmique de ce corps en mouvement est fragmentée par le montage, plus le locuteur acquiert cette corporalité pleine qui se fait le siège d'une voix ouvertement performantielle. La multiplication des plans, en exhibant le caractère construit de la représentation filmique, tend à abstraire le *hic et nunc* de l'énonciation verbale, à augmenter l'écart qui sépare le lieu profilmique de l'espace diégétique (comme l'illustrent de façon emblématique les chorégraphies de Busby Berkeley). C'est lorsque le filmage se soumet à la performance filmée que la valeur attractionnelle de la voix est la plus forte. La voix-attraction est d'autant plus marquée qu'elle est synchrone, c'est-à-dire qu'elle fait *corps* avec l'acteur qui la profère. C'est pourquoi elle est entravée par le principe du découpage, qui conduit presque inévitablement à rejeter provisoirement le locuteur hors-champ.

La construction d'un espace diégétique par le montage n'oblitére toutefois pas toute possibilité de manifestation de la voix-attraction, dans la mesure où cet espace peut prendre le locuteur comme centre de gravité. La voix ne s'adresse dès lors plus directement au spectateur du film, mais à son « homologue » situé à l'intérieur de l'univers diégétique. Le sujet parlant (ou chantant) est dans ce cas l'objet du regard d'autres personnages, sa voix interagissant avec d'autres éléments diégétiques. Cette situation intermédiaire constitue une modalité d'intégration de la voix qui tend à diminuer sa valeur attractionnelle. Lorsqu'il s'agit d'un spectacle dans le film, le degré d'attraction est proportionnel à l'autonomisation du spectacle. Si les paroles adressées par le locuteur aux spectateurs diégétiques ne font pas totalement partie de son spectacle, elles relèvent dans l'ensemble, en dépit de leur valeur attractionnelle, de la voix-action. La mise en scène d'une interaction verbale entre l'espace de la performance et l'espace spectatoriel représente néanmoins une possibilité d'intrusion d'un régime vocal plus fortement attractionnel. Le film recrée une situation de transmission orale telle que le cinéma l'avait connue à l'époque du

boniment, mais à l'intérieur même de son univers diégétique. La voix ne peut donc être considérée indépendamment de l'espace dans lequel elle se déploie: espace de la salle pour le bonimenteur, espace scénique pour l'acteur de théâtre, espace diégétique pour les personnages de films parlants.

#### IV.4 PERPETUATION DES ATTRACTIONS SOUS FORME VOCALE

Dans la production hollywoodienne «classique», l'intégration de parties chantées touche la plupart des genres de façon ponctuelle (nombreux sont les westerns donnant lieu, au saloon ou autour d'un feu, à une séquence de chant, ainsi que les films noirs dont l'actrice séductrice apparaît sur la scène d'un night-club). De même, le principe du leitmotiv qui intègre à la musique extradiégétique la mélodie d'un air chanté préalablement par un personnage est une pratique courante dont la chanson «As Time Goes By» de *Casablanca* (Michael Curtiz, 1942) est un exemple emblématique. Toutefois, ces moments chantés n'instaurent un régime vocal dominant (à l'échelle du film) que dans certaines comédies musicales. Sinon, le chant n'occupe pas une place centrale, mais fait en quelque sorte office de digression. Selon les cultures, ce rapport entre la voix parlée et la voix chantée peut s'inverser, le chant s'imposant par exemple comme une norme dans le cinéma indien où, pour la période 1931-1954, on ne compte que deux films dépourvus de chansons<sup>420</sup>. Le genre du *musical*, unique exemple cité par Gunning pour illustrer la persistance du «cinéma des attractions» (les pratiques dites d'«avant-garde» exceptées), présentent une série de numéros chantés ou dansés qui constituent des entités discrètes passablement autonomes, à l'instar des diverses «attractions» dont étaient composées les séances de «cinéma» au début du XX<sup>e</sup> siècle, ainsi que divers spectacles populaires du siècle précédent. Il me paraît donc intéressant de commenter la façon dont ce genre a été théorisé sur le plan narratologique.

Martin Rubin inscrit la comédie musicale – et particulièrement les films de Busby Berkeley – dans ce qu'il nomme la «tradition du spectacle» (comprenant le vaudeville, le *minstrel*, le burlesque, etc.), définie par une volonté de créer la variété et l'étonnement au

420. Selon Manjunath Pendakur, *Indian Popular Cinema*, Cresskill: Hampton Press, 2003, p. 123.

lieu de privilégier la continuité et l'unité propres au cinéma « classique ». Bien qu'il conçoive notamment le *musical* comme un héritage des attractions du musée Barnum<sup>421</sup>, Rubin n'utilise pas le terme d'« attraction » (ni le contexte théorique afférent) pour qualifier les moments chantés. Il en va également ainsi des principaux théoriciens du cinéma qui ont traité de ce genre. La terminologie couramment utilisée est celle de l'opposition entre narration et spectacle, problématisée à travers la notion d'« intégration ». Rubin s'attaque à la conception qui prévaut chez les historiens du cinéma selon laquelle la comédie musicale se serait progressivement développée dans le sens d'une plus grande intégration des numéros chantés dans le récit. Il est vrai, en effet, qu'un film comme *Le Chanteur de jazz* (Alan Crosland, 1927) – qui n'est certes pas une « comédie » musicale, mais qui allie narration et séquences chantées – dément ce postulat évolutif (sous-tendu par une volonté de valoriser la cohérence des œuvres) en recourant à diverses stratégies d'inscription des performances chantées dans un réseau de regards diégétiques qui assure la constance des enjeux narratifs. Rubin propose à juste titre d'admettre une tension constante entre intégration et non-intégration, nécessaire car véritablement fondatrice du genre<sup>422</sup>. Cette bipolarité permet à Rubin de ne pas faire de la narration le critère absolu de l'évaluation des effets des moments chantés. Il échappe ainsi à la normativité d'Alain Masson, pour qui un *musical* « est un film dont la narration se trouve interrompue ou continuée sous forme de chant et parfois de danse »<sup>423</sup>. Dans la définition de Masson, la narration constitue en effet le point de repère à partir duquel on peut estimer la rupture provoquée par l'intrusion d'un moment spectaculaire. L'auteur envisage deux modalités : le prolongement de la narration à travers le chant, qui s'intègre à l'ensemble en présentant des éléments narratifs, ou l'interruption du récit, dont il importe alors que la « qualité même soit mise en valeur ». Cette opposition correspond à ce que nous avons dit du couple attraction/narration, même si Masson ne fait pas référence à ces notions.

Dans l'ouvrage qu'il consacre à la comédie musicale, Rick Altman ne recourt pas non plus à cette terminologie, mais se borne à qualifier les passages chantés ou dansés de « numéros ». Sa thèse principale

421. Martin Rubin, *Showstoppers. Busby Berkeley and the Tradition of Spectacle*, New York: Columbia University Press, 1993, p. 15.

422. *Ibid.*, pp. 12-13.

423. Alain Masson, *La Comédie musicale*, Paris: Stock, 1981, p. 32.

repose sur le concept de «structure bplex» qui lui permet de rendre compte de l'absence (ou du moins de l'extrême superficialité) des paramètres communément rattachés à la narration «classique» (linéarité, succession, causalité) au profit d'un principe de mise en parallèle de situations opposées. Cette structure consiste à présenter «deux foyers d'intérêt constitués par un couple de stars de sexe opposé et se réclamant de valeurs divergentes»<sup>424</sup>. Il faut souligner que ce modèle, bien que n'accordant pas le primat à la narration, tend néanmoins à offrir un principe unifiant qui enrayer l'énergie centripète de l'attraction. En ce sens, le jumelage sur lequel se fonde la structure bplex peut être considéré comme un facteur d'intégration, même si cette dernière n'est pas (ou peu) «narrative». Les numéros perdent en effet de leur autonomie en raison des liens qu'ils tissent à travers la construction en bplex. Toutefois, cette cohérence ne s'élabore pas avec une force comparable à celle de la mise en intrigue: alors que la chronologie détermine une place spécifique pour chaque élément narratif, les paires de représentations contrastées de la structure bplex ne s'inscrivent pas dans un ordre strict et sont soumises à un fort principe de répétition<sup>425</sup>. En outre, les contrastes résultent d'une exhibition d'éléments visuels ou sonores, et donc participent de l'esthétique attractionnelle. Décors, gestes, musique sont affichés en tant que tels, dans leur dimension spectaculaire pure. Dans ce contexte, la voix devient un facteur essentiel, ne serait-ce que parce qu'elle est un critère majeur de distinction entre le masculin et le féminin. Lorsqu'il aborde la question du «style personnel» des acteurs, Altman insiste sur l'importance des voix en mentionnant l'exemple du couple Maurice Chevalier/Jeannette Mc Donald, interprètes dont les caractéristiques vocales contrastent fortement<sup>426</sup>. Si les paroles des chansons nourrissent parfois le contexte narratif, la voix est également susceptible de produire du sens.

Le modèle d'analyse de l'intégration des numéros musicaux qui me semble le plus intéressant est celui qu'esquisse brièvement John Müller dans l'introduction d'un article sur Fred Astaire. Son entreprise classificatoire repose sur une conception graduelle et aboutit à six types distincts que Müller qualifie comme suit, selon un ordre croissant d'intégration :

424. R. Altman, *La Comédie musicale hollywoodienne*, op. cit., p. 30.

425. *Ibid.*, p. 44.

426. *Ibid.*, p. 56.

«1. Numéros sans pertinence aucune pour l'intrigue. 2. Numéros qui contribuent à l'esprit ou au thème. 3. Numéros dont l'existence est pertinente pour l'intrigue, mais dont le contenu ne l'est point. 4. Numéros qui enrichissent l'intrigue mais qui ne la font pas avancer. 5. Numéros qui font avancer l'intrigue, mais pas par leur contenu. 6. Numéros qui font avancer l'intrigue par leur contenu.»<sup>427</sup>

Plutôt que de considérer ces catégories de façon rigide, il me semble plus judicieux de les associer à des cas prototypiques qui sont autant de bornes jalonnant un spectre continu, valable tant à l'échelle d'un film entier que pour un numéro donné. Les critères retenus par Müller pour discuter les diverses modalités de mise en relation de l'intrigue et des numéros ne sont toutefois pas stables, ce qui provoque des indécisions regrettables. Il me paraît par exemple difficile d'estimer ce que l'«esprit» ou le «thème» implique en termes de narrativité, alors qu'il s'agit du principal trait distinctif retenu par Müller pour écarter les formes non narratives. Cette question du «thème» devient d'autant plus floue lorsque Müller précise que la revue possède un degré d'intégration supérieur à celui du vaudeville parce que les numéros y sont parfois ancrés dans un «thème» commun<sup>428</sup>; or, tout comme le vaudeville, la revue n'est pas considérée par Müller comme une forme narrative. Les deux paramètres «intégration» et «narration» ne sont donc pas traités conjointement, alors que l'on pourrait considérer la seconde comme l'une des manifestations de la première. En fait, les catégories de Müller semblent partagées entre deux opérations différentes d'élaboration de la fiction cinématographique: d'une part la mise en place d'un univers diégétique qui consiste à définir les personnages et l'environnement, d'autre part la mise en intrigue qui inscrit ces éléments dans un développement temporel. Cette distinction permet selon moi de regrouper les types comprenant des numéros dépourvus de contenu narratif (comme les types 3 et 5, qui reposent en fait sur le même principe). Lorsque l'intégration se résume au statut fictionnel des personnages au sein de ce que Rick Altman a nommé la «comédie-spectacle», nous sommes en présence d'une *motivation diégétique*

427. John Müller, «Fred Astaire and the Integrated Musical», *Cinema Journal*, Vol. 24, N° 1, p. 30.

428. *Ibid.*, pp. 37-38, note 2.

(les personnages sont aptes à chanter et danser avec brio, ce qu'attestent leurs performances) qui existe indépendamment de tout facteur « chrono-logique ». L'élaboration de la diégèse implique également une dimension spatiale: c'est ce qu'aborde brièvement Müller dans une sorte de « sous-type » au statut ambigu lorsqu'il classe, parmi les numéros dénués de toute pertinence narrative, ceux « dans lesquels les personnages principaux de l'intrigue regardent la performance avec le public, mais n'y participent pas »<sup>429</sup>. Ce cas de figure, selon lui assez fréquent, soulève la question de l'interaction entre *chant* et *champ*. La création d'un espace spectatorial interne à la diégèse, où des personnages – qu'ils soient principaux, secondaires ou de simples figurants – prennent le relais du regard du spectateur, constitue un facteur déterminant d'intégration. Cette configuration ne ressortit cependant à la narration que si les spectateurs diégétiques endossent une fonction actancielle donnée, ce qui est par exemple le cas du personnage de Yudelson au début du *Chanteur de jazz* durant la chanson « Ragtime Jackie ». Dans ce cas, la voix ne s'arrête pas à la rampe de la scène, mais investit l'espace, soit directement (elle atteint le bar adjacent où se trouve l'orthodoxe Yudelson), soit par le relais de ce protagoniste qui se rend chez le père du jeune chanteur pour lui rapporter ce dont il a été le témoin; ainsi des espaces physiques (mais aussi sociaux, culturels, etc.) présentés comme disjoints au niveau de la topographie diégétique se voient-ils interconnectés par un travail aux implications « narrativisantes » sur la mise en scène de la réception d'une performance vocale. C'est pourquoi la narrativité n'est pas forcément propre au numéro lui-même, mais aux effets qu'il provoque sur ceux qui y assistent.

Il faut par ailleurs prendre en compte le degré de narrativité que présentent les numéros en eux-mêmes, car chacun d'eux peut également se concevoir entre les deux pôles attraction/narration. L'intégration s'opère alors avec d'autant plus de prégnance si le contenu narratif du numéro « enchâssé » fait écho au récit enchâssant en le nourrissant d'éléments nouveaux. La répétition sous une forme chantée d'une information narrative déjà transmise par le dialogue ou en actes participe certes de l'intégration – la « redondance » constitue un trait stylistique fondamental de la pratique « classique » – mais le degré d'intégration narrative est supérieur

429. *Ibid.*, p. 28.

lorsque le numéro prend directement place dans la chaîne de l'organisation syntagmatique du récit enchâssant, lorsqu'il occupe une position nodale dans le développement de l'intrigue. On peut donc faire l'hypothèse que le degré d'intégration des numéros dans le film varie en fonction de leur participation aux processus de diégétisation (qui correspondent, selon les aspects considérés, aux cinq premiers types de Müller), et à la transmission d'informations narratives qui, du point de vue de la voix, s'effectue soit sur la base de la *narrativité interne* des paroles proférées dans le chant, soit par le biais de la *narrativité externe* de la situation qui offre son cadre à la performance chantée. Selon Roger Odin, les processus de diégétisation et de narrativisation contribuent au phénomène global de fictionalisation<sup>430</sup>. Or, dans la comédie musicale, l'un des indices de l'intégration dépend du statut que l'on attribue, en termes de fictionalité, à l'auteur de la performance: lorsque l'intégration narrative est conséquente, l'acteur en tant que personne réelle – connue du public pour ses performances dans un contexte différent (scénique, télévisuel, etc.) – s'efface au profit du personnage du film.

Moins les numéros des comédies musicales donnent lieu à des processus de diégétisation/narrativisation, plus leur valeur attractionnelle et leur parenté avec le cinéma des premiers temps sont grandes. Comme le cinéma bonimenté, le *musical* est un genre où l'adresse au spectateur du film n'est pas nécessairement tabou. La réflexivité qui le distingue des autres genres hollywoodiens vise, comme l'a montré Jane Feuer, à simuler une relation directe avec le public, à créer cette immédiateté qui est le propre d'un théâtre populaire – mais aussi, comme on l'a vu, du «cinéma» *parlé* – auquel est associé un fort esprit communautaire<sup>431</sup>. Même si cette référence relève du «mythe» pour Feuer (l'institution cinématographique est, on le sait, dépourvue de toute dimension interactive), elle n'en demeure pas moins la source d'une pratique attractionnelle (située au *second niveau* d'intégration, bien que prétendant illusoirement relever du premier) qui exploite dans une large mesure la voix.

Il n'est pas fortuit que la parole se manifeste dans la plupart des moments d'attraction sous la forme du chant, mode d'expression qui permet d'exploiter au maximum les capacités (rythmiques,

430. Roger Odin, *De la fiction*, Bruxelles: De Boeck, 2000.

431. Jane Feuer, «The Self-Reflective Musical and the Myth of Entertainment», in Steven Cohan (dir.), *Hollywood Musicals. The Film Reader*, London; New York: Routledge, 2002.



mélodiques, expressives) de la voix indépendamment du contenu linguistique. Paul Zumthor a souligné à plusieurs reprises les vertus orales propres au chant. Pour lui, la force particulière du chant influe sur ce qu'il appelle le « mode » de la performance :

« Dans le *dit*, la présence physique du locuteur s'atténue plus ou moins, tend à se fondre parmi les circonstances. Dans le chant, elle s'affirme, revendique la totalité de son espace. C'est pourquoi la plupart des performances poétiques, dans toutes les civilisations, ont toujours été chantées: et pourquoi, dans le monde d'aujourd'hui, la chanson, en dépit de son avilissement par le commerce, constitue la seule véritable poésie de masse. »<sup>432</sup>

Le *dit*, proféré par un acteur visualisé, est du côté de la voix-action, qui se fait événement et interagit directement avec les « circonstances » (diégétiques). Le chant se soustrait par contre à cette immersion tout en s'affirmant comme présence. La description de Zumthor correspond bien, dans notre typologie des régimes vocaux, à la voix-attraction. Généralement, le chant procède d'une exploitation marquée de la fonction émotive du langage, et constitue le lieu privilégié d'accentuation du pôle du destinataire (fonctions phatique et conative). Zumthor voit dans la chanson un engagement du collectif, notamment parce que l'auditeur est plus intimement pris à partie par le locuteur, qui l'enjoint à devenir lui aussi « auteur » de la performance. Cette participation se voit naturellement renforcée si l'auditeur connaît également les paroles du chant. Une telle exacerbation de la « relation de proximité » avec le public s'observe de façon manifeste dans ce que Rick Altman appelle la « comédie-folklore », sous-genre de la comédie musicale qu'il définit entre autres par son ancrage dans une certaine réalité quotidienne (opposée au monde du *show business* par exemple) et par le recours à des airs populaires. En véhiculant principalement des valeurs communautaires, la comédie-folklore exacerbe nécessairement sa relation au spectateur. Ainsi Altman dit-il de la première réalisation sonore de King Vidor (*Hallelujah*, 1929), « comédie-folklore » qui propose elle-même une représentation collective du chant comme l'expression de « l'âme » d'un peuple (le titre choisi pour l'exploitation française fut significativement *L'Âme*

432. P. Zumthor, *Introduction...*, *op. cit.*, p. 178.

*noire*): «Le public de *Hallelujah* retrouva nombre de ses vieux airs favoris – des *negro spirituals* essentiellement – si bien qu'à chaque séance la salle ne restait pas muette»<sup>433</sup>. Le chant de la «comédie-folklore» n'en demeure pas moins un facteur essentiel de l'élaboration de la diégèse en ce qu'elle exprime les états d'âme des personnages. Elle constitue à cet égard un instrument d'intériorisation comparable aux informants présents dans certaines situations parlées (confession, aparté) plus proches de la voix-action. Altman illustre cette fonction du chant en disant de deux acteurs de *musicals*, Judy Garland et Frank Sinatra, qu'ils «semblent chanter sans faire le moindre effort, sans entraînement, d'une façon qui n'évoque ni l'opéra, ni la scène de concert mais qui se donne comme l'expression d'une émotion personnelle»<sup>434</sup>. Une telle démarche naturalisante s'accompagne inévitablement d'une atténuation de la valeur attractionnelle du chant. Lorsque la voix-attraction est au contraire exhibée, elle se nourrit de l'intermédialité pour référer, dans le film, à un contexte spectaculaire qui repose majoritairement sur la voix vive.

Comme le boniment du «muet», le chant *captive*, transmet son énergie aux spectateurs. En désaliénant la parole de sa stricte soumission au sens, le chant acquiert une certaine «gratuité» qui est le propre de l'agencement libre de l'esthétique des attractions. Ce détachement des contingences diégétiques se manifeste concrètement dans les *musicals* par le recours quasi généralisé à la postsynchronisation, qui donne fréquemment lieu à un synchronisme large. Du fait de ses autres modes de consommation (sur disque, sur les planches de Broadway, à la radio, etc.), le chant du film n'appartient d'ailleurs pas totalement aux personnages qui sont les supports de sa profération, fussent-ils incarnés par les chanteurs réels. Les acteurs se font les porteurs d'une parole qui, parce qu'elle les dépasse, appartient également au spectateur. Cette dimension fondamentalement collective du chant est exploitée par René Clair dans ses réalisations du début des années 30, et par Alain Resnais dans *On connaît la chanson* (1997), film dont le titre à double sens renvoie au principe de reconnaissance (du Même) qui est à la base d'un sentiment communautaire (le sujet est un «on»). Personnages et spectateurs

433. R. Altman, *La Comédie musicale hollywoodienne*, op. cit., p. 344.

434. *Ibid.*, p. 308.

partagent des expériences similaires. C'est pourquoi il importe que les situations chantées de la trilogie parisienne de Clair ou d'*On connaît la chanson* soient ancrées dans le quotidien des protagonistes: loin des échappées dans le faste qui caractérisent les comédies musicales, ces films demeurent proches de l'environnement de tous les jours de leurs spectateurs. Un film comme *Tout l'or du ciel* (*Pennies from Heaven*, Herbert Ross, 1981) pourrait sembler obéir au même principe qu'*On connaît la chanson*, puisque des voix chantées surgissent qui n'appartiennent pas aux acteurs (dont les lèvres se meuvent pourtant de façon synchrone). Pourtant, il s'en distingue radicalement dans la mesure où il ne met pas en œuvre cette fonction sociale du chant à travers une interpellation de l'audiospectateur. En effet, non seulement le film de Ross est une reconstitution (son intrigue se déroule en 1932) mais le décor, subitement affecté d'un degré supérieur d'irréalité, est l'objet de transformations lorsque les acteurs commencent à chanter. Tandis que l'espace diégétique s'y déploie pour permettre les évolutions chorégraphiques, les personnages de Resnais ne renvoient aucunement par leurs attitudes à la présence de la chanson.

La dominante attractionnelle de la parole chantée n'interdit pas toute situation *over*. Le chant *over* interpelle plus fortement l'audiospectateur que la voix-over parlée; on ne s'étonnera donc pas de le trouver plutôt en début de film, comme dans *Brigadoon* (Vincente Minnelli, 1954), où la fonction narrative subsiste (le chœur puis un chanteur solo dépeignent la situation de deux chasseurs perdus dans une forêt), alors que la dimension chantée introduit une situation quelque peu décalée, presque onirique. Le chant *over* est généralement utilisé pour renvoyer à l'oralité des récits mythiques: il en va ainsi des chants liturgiques de *La Couleur de la grenade* (Sergei Paradjanov, 1968), film quasiment dépourvu de sons synchrones qui nous conte l'histoire d'un troubadour arménien, ou du récent *Chant de la fidèle Chunhyang* (Im Kwon Taek, 2000), où une voix-over parlée alterne avec le Pansori, chant narratif traditionnel coréen dont l'interprète nous est présenté sur scène dès la séquence liminaire avant de se manifester *over* en mêlant informations narratives et évocation poétique. Ces films mettent en scène des figures qui s'apparentent au conteur médiéval dont la performance constitue la base de la théorie générale de l'oralité que propose Paul Zumthor.

#### IV.5 L'INTEGRATION MUSICALE DE LA PAROLE

La conception graduelle qui régit l'opposition attraction *vs.* intégration s'applique par ailleurs aux différents types d'interventions chantées. S'il me paraît judicieux d'affirmer que le chant introduit nécessairement une composante attractionnelle – et ce d'autant plus que le contenu des paroles est faiblement narratif – il faut en outre s'interroger sur le caractère graduel du chant lui-même. Il arrive en effet, dans certaines comédies musicales, et notamment dans celles où intervient Fred Astaire (cet artiste s'étant fait une spécialité de l'insertion progressive du chant ou de la danse), que la voix parlée se transforme graduellement en une voix chantée. Ce phénomène, que je propose de nommer *intégration musicale (du langage parlé dans le chant)*, obéit au principe de continuité du « fondu sonore » qui, d'après Rick Altman, permet d'atténuer les « sautes » résultant du surgissement de la musique extradiégétique et de la suppression corrélative des bruits diégétiques.

Comme exemple de pont sonore, Altman cite le morceau chanté « Fancy Free » de *Top Hat (Le Chanteur du dessus)*, Mark Sandrich, 1935), film dans lequel Fred Astaire passe de façon imperceptible de la conversation à la chanson, « simplement en rythmant et en modulant les accents de sa voix »<sup>435</sup>. La parole chantée impliquant en elle-même une composante attractionnelle, l'intégration musicale de la voix peut être considérée comme inversement proportionnelle à l'intégration narrative de la performance vocale, à moins que les circonstances proprement dites du chant ne nourrissent le récit (cas de « narrativité externe »), comme cela se produit dans *Top Hat* où les claquettes du chanteur, empêchant le personnage féminin interprété par Ginger Rogers de dormir à l'étage inférieur, sont à l'origine de la rencontre du couple. Dès son entame qui correspond à l'ouverture de la porte de la chambre d'hôtel, futur espace « scénique », cette séquence de *Top Hat* comprend une musique extradiégétique au niveau sonore élevé (par rapport aux standards des scènes dialoguées) qui, dans l'esprit du spectateur habitué à ce genre de films, « appelle » la performance d'Astaire, alors que celui-ci ne se met à chanter que trois minutes plus tard. Cette musique est d'autant plus perceptible que la séquence précédente au « London Trackeray Club » se déroulait dans un lieu où un silence total était

435. *Ibid.*, pp. 81-82.

exigé. Le contraste entre le silence pesant du club et la musique extradiégétique de cette séquence anticipe l'autonomisation de la performance chantée et dansée d'Astaire. Notons par ailleurs que le faible degré de perceptibilité de la transition du langage parlé vers le chant tient également au jeu d'Astaire, dans un premier temps nullement affecté par ce changement : lorsque l'acteur commence à chanter, il tient sa cigarette à la main, reste assis sur l'accoudoir du sofa puis se sert machinalement un verre.

Les modalités d'intégration musicale de la parole sont nombreuses. Ainsi Charles Wolfe propose-t-il de considérer deux pôles pour rendre compte des moments chantés dans *Le Chanteur de jazz*, la « performance vocale » et le « drame vocalisé »<sup>436</sup>, qui correspondent à ce que j'ai proposé d'appeler la voix-attraction et la voix-action. André Wyss évoque « un véritable *continuum*, représenté par les gradations de nuances de la voix parlée-chantée et de la voix parlée-rythmée »<sup>437</sup>. On peut néanmoins définir quelques jalons sur la base de certains traits spécifiques. Les critères de distinction entre voix parlée et voix chantée relèvent de deux catégories : la ligne verticale des tons, et la ligne horizontale du « rythme » organisée en mesures et en temps. La première permet par exemple de différencier le chant de la mélodie parlée, car celui-là est fondé sur des écarts importants et mesurables au niveau de la hauteur des tons. Quant à la mélodie parlée, elle se distingue de la parole par son organisation horizontale, tout comme le chant. Dans le langage parlé, les écarts instaurés par la ligne verticale sont relativement faibles. Des variations s'observent néanmoins d'un individu à l'autre (chacun actualisant les phonèmes à sa manière, introduisant une courbe mélodique qui lui est propre, frappant certaines syllabes d'un accent particulier), d'un texte à l'autre (en fonction de son degré de « poéticité ») et d'une langue à l'autre (certaines langues sont plus « mélodiques » que d'autres). Tous ces paramètres influent sur la perception d'une manifestation vocale située entre le parlé et le chanté. Les marges de réalisations que la langue offre au locuteur et qui permettent la constitution de ce qu'Ivan Fónagy appelle le « style vocal » s'appliquent notamment au degré de mélodicité, c'est-à-dire à la présence de régularités sur la courbe qui se dessine entre l'axe temporel et celui de la hauteur des sons.

436. Ch. Wolfe, « Vitaphone Shorts... », *art. cit.*, p. 69.

437. André Wyss, *Eloge du phrasé*, Paris : PUF, 1999, p. 154.

Comme le constate Fónagy, la musicalité paraît être directement proportionnelle à la redondance de la distribution des fréquences sur la ligne horizontale du temps, c'est-à-dire à la « rythmicité » des éléments phonétiques<sup>438</sup>. La redondance dépend en partie du texte qui est proféré puisque certaines « figures » poétiques (rimes, anaphores, assonances, allitérations, répétitions, etc.) peuvent la renforcer, mais elle résulte aussi du style vocal du locuteur.

Dès lors que l'on refuse l'idée d'un partage net entre parlé et chanté, on peut s'interroger sur des démarches artistiques qui consistent à décliner certaines possibilités intermédiaires. C'est ce dont discute Alban Berg dans l'introduction à sa partition de *Lulu* (1935), opéra dans lequel certains passages du texte de Wedekind sont simplement parlés en l'absence de musique. Exposant une conception graduelle des phénomènes vocaux, Berg définit six cas de figure compris dans un éventail d'alternatives. Selon Étienne Barilier, il s'agit pour le compositeur autrichien de « réaliser la transition entre le théâtre et l'opéra [...], c'est-à-dire entre parole et chant »<sup>439</sup>. L'insertion plus ou moins appuyée d'une composante rythmique ou mélodique permet les « fondus sonores » évoqués par Altman à propos de la comédie musicale.

Le spectre couvert tant par l'intégration musicale de la parole que par l'intégration narrative des occurrences vocales est extrêmement large et donne lieu à un si grand nombre de possibilités qu'il semble plus pertinent d'opérer des analyses individuelles plutôt que de proposer une typologie prétendument exhaustive. À cet égard, la conception graduelle que j'ai proposée pour différents concepts me paraît être le modèle le plus adéquat.

#### **IV.6 LA VOIX-ATTRACTION PARLÉE : L'EXEMPLE DES MARX BROTHERS**

Le champ d'application de la conception bipolaire attraction *vs.* intégration ne se réduit pas au seul genre de la comédie musicale, ou aux séquences chantées de films qui, à l'instar d'œuvres récentes comme *Jeanne et le garçon formidable* (Olivier Ducastel et Jacques Martineau, 1998), *8 Femmes* (François Ozon, 2002) ou *Pas sur la bouche* (Alain Resnais, 2003), ont été réalisées dans la tradition

438. I. Fónagy, *op. cit.*, p. 310.

439. Étienne Barilier, *Alban Berg*, Lausanne: L'Âge d'Homme, 1992, p. 248.

française du film «chanté» initiée par René Clair et perpétuée notamment par Jacques Demy. L'insertion de monologues ou de dialogues comiques dans un film, redevable de pratiques scéniques de la fin du XIX<sup>e</sup> siècle telles que les «*minstrel shows*»<sup>440</sup>, tend également à accroître la valeur attractionnelle du régime vocal dominant.

L'un des exemples les plus typiques d'une esthétique de la voix-attraction qui ne repose pas exclusivement sur le chant est celui des films des Marx Brothers. Issue de l'exploitation foraine, du vaudeville puis de la scène de Broadway, cette famille d'artistes reconduit effectivement des modes de représentation conçus pour des spectacles vivants, où la parole vive et les performances musicales occupent une place «structurelle». Il est intéressant de noter que l'apparition remarquée des Marx dans l'industrie cinématographique est strictement contemporaine de l'«*inter-règne*» (1929-1934). Le discours historiographique portant sur leur filmographie suggère d'ailleurs souvent un déclin qui se serait manifesté à partir d'*Une nuit à l'opéra* (*A Night at the Opera*, 1935)<sup>441</sup>. Or ce film correspond chez les Marx au passage de la Paramount à la MGM, fréquemment associé par les commentateurs à une certaine domestication de leurs films, en dépit de la constante résistance de ces artistes aux propositions des scénaristes et donc, d'une certaine manière, à l'intégration narrative<sup>442</sup>. L'adresse verbale au public utilisée par Groucho dès leur premier film, relique manifeste de ses propres performances scéniques – un procédé auquel le producteur Monta Bell se serait vivement opposé<sup>443</sup> – constitue peut-être, en ce qui concerne la parole, l'élément le plus manifeste de cet emprunt à un mode de présentation fortement attractionnel. D'ailleurs, leurs deux premiers films sonores, réalisés dans les années d'institutionnalisation du cinéma parlant, *Noix de coco* (*Cocoanuts*, 1929) et *L'Explorateur en folie* (*Animal Crackers*, 1930), sont des adaptations de pièces qui furent d'abord montées à Broadway.

Les films des Marx accordent une place centrale à la parole; c'est pourquoi, pour Benayoun, ils sont à compter parmi ceux qui «ne tolèrent pas le doublage»<sup>444</sup>. Les jeux de mots et les monologues

440. R. Altman, *La Comédie musicale hollywoodienne*, op. cit., p. 221.

441. Voir par exemple Georges Sadoul, *Histoire du cinéma mondial*, op. cit., p. 289, ou Jean-Louis Passek, *Dictionnaire du cinéma*, Paris: Larousse, 1995, p. 1421.

442. Voir Robert Benayoun, *Les Frères Marx*, Paris: Seghers, 1980, pp. 36-38.

443. Selon Allen Eyles, *The Complete Films of the Marx Brothers*, New York; Secaucus: Citadel Press, 1992, p. 31.

444. R. Benayoun, *Les Frères Marx*, op. cit., p. 25.

déclamés avec un débit frénétique se prêtent en effet mal à la traduction. Ce constat est en outre révélateur du fait que les répliques volubiles de Groucho – un acteur qui, à l’instar de Sacha Guitry, laisse peu parler les autres – valent pour elles-mêmes, pour leurs caractéristiques langagières et phonétiques. Cette prédominance des fonctions poétique et métalinguistique fait de la prise de parole une manifestation singulière et autonome, une « attraction ».

Afin d’examiner ce type de manifestations vocales, je propose d’examiner la première apparition d’un des frères Marx dans un film parlant, celle de Groucho au début de *Cocoanuts*. Après une séquence d’introduction qui, accompagnée d’une musique mais dépourvue de parole synchrone, nous a présenté la plage et ses baigneuses qui s’adonnent collectivement et en rythme à des exercices de gymnastique quasi chorégraphiés, un gros plan sur une enseignante nous indique que nous pénétrons dans « l’Hôtel de Cocoanut ». Dès le début de cette séquence filmée en un seul plan fixe (excepté quelques recadrages presque imperceptibles et un bref panoramique à la toute fin), M. Hammer (Groucho) se trouve déjà sur les marches du grand escalier de l’hôtel. Alors qu’il s’apprête à descendre, une cohue de grooms appartenant à cet établissement qu’il dirige le rejoignent et exigent qu’il leur paie leurs gages. Le patron se lance alors dans un discours aussi prolixe qu’illogique visant à convaincre les employés qu’il leur faut renoncer à leur salaire pour gagner leur liberté. Dans cette scène, tout concourt à mettre en évidence l’acte de parole de Groucho : il s’adresse à un auditoire collectif non individualisé (« Nous voulons être payés » est l’unique phrase articulée que les grooms prononcent en chœur)<sup>445</sup>, monopolise la parole et occupe une position centrale dans le plan, surplombant dans son veston foncé le groupe à casquette blanche. Ce public diégétique contribue, en décrivant un demi-cercle autour de lui, à naturaliser la situation d’énonciation, mais celle-ci rappelle néanmoins fortement les monologues scéniques. En effet, lorsqu’un serveur lui apporte deux télégrammes, M. Hammer les lit successivement à haute voix à

445. Précisons à propos de cette représentation d’une masse uniforme que les rôles des « garçons » d’hôtel sont significativement tenus par des femmes qui, ultérieurement, jouent le rôle de danseuses lors d’une « marche » sur laquelle se termine la séquence. Il s’agit là de *girls* comme celles dont il est question chez Siegfried Kracauer lorsqu’il examine la standardisation résultant de l’industrialisation de la culture dans le courant des années 1920-1930. Le sociologue remarquait par exemple que « ces produits des usines de distraction américaines ne sont plus des jeunes filles individuelles, mais des groupes indissolubles [...] dont les mouvements sont des démonstrations mathématiques » (*Figures de ville...*, *op. cit.*, p. 70).



l'assemblée, utilisant leur contenu pour émettre des commentaires décalés. L'absence de travail au niveau du montage témoigne également d'une volonté de restituer l'intégralité de la performance de l'acteur, qui est probablement en partie improvisée.

On ne peut pas dire que ce passage ne soit pas intégré au récit du film. Toutefois, si quelques éléments narratifs sont présents (mauvaise situation financière du patron, annonce de l'arrivée probable de personnages importants), ils sont noyés dans la logorrhée « martelée » par le justement nommé « Hammer ». Sa verve perturbe totalement le statut de mise en place de la diégèse que devrait endosser cette première séquence dialoguée. La suite montre combien ce passage obéit à un mode d'agencement qui relève essentiellement de la juxtaposition : sans souci de raccord, nous voyons les grooms effectuer ensemble des pas rythmés, alors qu'une musique extradiégétique se fait tout à coup entendre. Ce bref moment musical acquiert le statut d'intermède entre deux séquences qui font intervenir les stars du film. L'enchaînement repose donc sur la succession de numéros, un principe qui régit l'ensemble du film. La parole constitue l'une des attractions de *Cocoanuts*, au même titre que la danse, le chant ou la musique instrumentale jouée devant la caméra.

Ces différents moments de spectacle ne sont néanmoins pas totalement indépendants du contexte diégétique, celui du boom immobilier de la Floride, omniprésent dans les déclarations de Groucho. Ainsi, lorsque l'intermède dansé des girls est terminé, Groucho fait une brève allusion à cette performance, ce qui confirme qu'elle a bien « existé » pour les autres personnages et n'a pas seulement été adressée au spectateur du film. Le premier long métrage des Marx multiplie d'ailleurs les situations de spectacles diégétisés. Groucho se fait par exemple bonisseur lors d'une vente aux enchères, ou s'improvise maître de cérémonie au mariage de Polly. L'intégration de ces numéros parlés s'effectue sur le mode de la diégétisation ; la narration occupe par contre une place secondaire, réduite à l'intrigue policière tenue dans laquelle s'imbrique l'histoire d'amour de Polly. Nous ne sommes donc pas totalement en présence de l'esthétique fragmentaire des spectacles de revue. Cette orientation fut d'ailleurs prise dès la réalisation de la version scénique de *Cocoanuts*, puisque les Marx s'opposèrent au projet initial du producteur Sam H. Harris de créer un show sur le modèle de la revue et demandèrent une pièce écrite<sup>446</sup>.

446. A. Eyles, *op. cit.*, p. 30.

Les caractéristiques que j'ai relevées dans la séquence d'ouverture de *Cocoanuts* s'observent dans la plupart des films des Marx, qui procèdent le plus souvent à une juxtaposition passablement aléatoire de moments de spectacle insérés dans une narration lâche. Ces séquences d'attractions sont basées sur la pantomime burlesque (actions frénétiques de Harpo, en quelque sorte « bonimentées » ou « doublées » verbalement par Chico), sur la musique (le piano de Chico, la harpe de Harpo), sur la chanson (Groucho) ou sur des sketches (les apartés ou monologues de Groucho, les duos de ce dernier et Chico). Chacun de ces personnages types se définit d'ailleurs par la façon dont il s'approprie la voix et la parole.

Ce que les Marx mettent en œuvre presque systématiquement dans leurs films – qui, bien que réalisés par d'autres, n'en sont pas moins avant tout marqués par leurs performances profilmiques – concerne ponctuellement d'autres cinéastes. Il en va par exemple ainsi de l'insertion d'un numéro comique dans un film où l'humoriste n'apparaît à aucun autre moment, à l'instar de l'intervention inopportune de Darry Cowl dans la salle du tribunal de *Assassins et voleurs* (Sacha Guitry, 1956) ou du sketch de Raymond Devos dans *Pierrot le fou* (Jean-Luc Godard, 1965). Dans ces deux cas, « l'entrée en scène » du comique introduit une stase narrative, puisque le contenu de ses paroles n'entretient aucun lien direct avec l'intrigue du film. La reconnaissance d'une star de la scène enraye par ailleurs la construction du personnage.

#### **IV.7 SYNTHÈSE DES TRAITS DEFINITOIRES DE LA VOIX-ATTRACTION ET DE LA VOIX-NARRATION**

Sur la base des considérations théoriques faites jusqu'ici, j'aimerais proposer une synthèse des paramètres sur lesquels repose la bipolarité entre voix-attraction et voix-narration. Comme on l'a vu, ces deux notions s'appliquent a priori à des phénomènes situés sur des plans distincts : les attractions concernent avant tout un certain rapport au public, alors que la narration se rapporte majoritairement à la façon dont est organisé le « texte ». Toutefois, les dimensions pragmatique et structurale ne sont pas dissociables : l'esthétique des attractions se manifeste également au niveau de l'agencement des unités (primat de la juxtaposition, autonomisation des « actes »), alors que le récit filmique ne peut être conçu indépendamment de l'activité « configurationnelle » du spectateur

(au sens de Paul Ricoeur), elle-même conditionnée par les stratégies de communication mises en œuvre par le film (ou plus largement par le dispositif). Alors que l'attraction agit au niveau des *stimuli*, la narration enjoint le spectateur à convoquer des schèmes de perception de la temporalité. Considérées comme de grands modes d'interaction entre la représentation et le public, ces deux catégories me paraissent conserver leur pertinence pour l'analyse des spectacles audiovisuels. Il faut néanmoins concevoir la notion de « narration » au sens large, c'est-à-dire plutôt dans les effets qu'elle produit que dans les caractéristiques spécifiques de l'organisation qu'elle propose. Le pôle de la voix-narration se définit en effet essentiellement par sa composante *intégrative*; c'est pourquoi d'autres types de mises en forme relèvent également de ce pôle.

Comparativement aux circonstances *live*, la voix-attraction du cinéma parlant ne peut proposer, en raison de la nécessaire intégration du son au support filmique, qu'une valeur pseudo-attractionnelle. Plutôt que de mobiliser le contexte d'une profération effective de la voix-attraction, certains films parlants *font référence* au mode d'interaction avec le public qui régit les spectacles vivants, incitant le spectateur à mobiliser des connaissances qu'il a acquises à leur propos, que cela soit par son expérience personnelle ou par des sources de diverses natures (parmi lesquelles on compte le cinéma, qui nourrit au même titre que d'autres productions significatives la représentation que l'on se fait d'autres pratiques). Dans l'ouvrage où il synthétise les principales théories cognitivistes appliquées au cinéma, Laurent Jullier souligne l'importance des rapprochements qu'effectue le spectateur pour reconstruire mentalement l'espace supposé de la source sonore, compréhension qui exige d'avoir « intégré des dispositifs scéniques comme ceux de l'opéra, du cirque, du spectacle de lanterne magique ou de la conférence »<sup>447</sup>. Si les deux pôles de la voix-attraction et de la voix-narration sont si instinctivement identifiables, c'est parce que le spectateur a assimilé les contextes culturels auxquels ils renvoient : la pièce de théâtre ou le numéro de music-hall pour l'une, le roman ou le théâtre légitimé pour l'autre.

Dans le régime de l'attraction, la voix cesse d'être soumise à la seule nécessité de véhiculer une signification et s'offre *en tant que telle*, dans les qualités propres à la performance orale. C'est pourquoi

447. Laurent Jullier, *Cinéma et cognition*, Paris : L'Harmattan, 2002, p. 105.

la présence physique du bonimenteur dans la salle confère à son intervention vocale une valeur attractionnelle nécessairement supérieure à celle dont disposent communément les locuteurs du cinéma parlant. En mettant entièrement son discours au service de la fonction référentielle, le bonimenteur pouvait néanmoins renforcer le pôle de la voix-narration. Pour le boniment comme pour la voix du parlant, les deux régimes s'avèrent possibles. Ces usages sont clairement distingués par W. Stephen Bush, un conférencier qui « théorise » sa pratique dans un article du *Motion Picture World* du 28 août 1908 en donnant quelques conseils à ses confrères :

« Ne criez jamais, et évitez la faute la plus fatale du style de l'aboyeur : « Ici on voit sur la gauche. » « Dans cette image nous apercevons la belle héroïne qui entre sur la gauche. » Ce style contrefait l'ancienne manière de l'homme qui décrit les charmes du squelette vivant et de la femme à barbe. Étudiez l'image avant d'essayer de faire une conférence à son propos. Donnez à la conférence la forme d'une narration cohérente qui, si elle était imprimée comme vous l'avez transmise, donnerait une bonne idée du sujet sans le secours des images. »<sup>448</sup>

Bush valorise ostensiblement la fonction de narrateur qui, à l'instar de la plupart des voix-over du parlant, renforce la cohérence du film. Par contre, il dénigre, pour des raisons évidentes de légitimation culturelle de son travail, le mode de l'adresse du « bonisseur ». Ce que Bush veut bannir relève précisément de la voix-attraction : pour lui, la référence à *l'image en tant que telle* (fonction réflexive du boniment) et à la *deixis* de l'énonciateur (les fonctions conative et phatique prenant le pas sur la fonction référentielle) sont à exclure au profit d'une narration impersonnelle qui permet de faire oublier la présence physique du bonimenteur. La comparaison avec le texte imprimé est symptomatique de l'appartenance de la voix-narration au domaine de *l'écrit*, alors que la voix-attraction constitue la pleine expression de *l'oralité*. Ces deux modes ont nécessairement coexisté, même si W. Stephen Bush suggère, en 1908, que l'aboyeur appartient à une période révolue de l'exploitation cinématographique. Son argumentation laisse au contraire penser que le régime de la voix-attraction est encore dominant chez un certain nombre de bonimenteurs.

448. Cité in R. Altman, *Silent Film Sound*, op. cit., p. 142.

On pourrait rétorquer que le bonimenteur qui commentait quelques années auparavant des films relevant du «cinéma des attractions» se pliait nécessairement à l'esthétique attractionnelle des bandes projetées. Toutefois, on a vu la liberté que certains bonimenteurs s'octroyaient par rapport à l'image. Altman cite l'exemple de conférences durant lesquelles Cecil Hepworth présentait les danses de Loïe Fuller (performativité du geste et exhibition de vêtements et lumières qui sont à tous égards emblématiques de ce qui constitue une «attraction» au sens où nous l'entendons ici) en les inscrivant dans le récit de Salomé, contant comment ce «personnage» usait de ses charmes pour séduire Hérode<sup>449</sup>. Le pianiste disposait d'une même alternative lorsqu'il accompagnait les films de chansons populaires: soit celles-ci étaient ancrées dans l'univers diégétique (leur existence étant motivée par la visualisation d'une source potentielle) et leurs paroles faisaient écho au contexte narratif du film, soit elles renvoyaient à l'espace même de la salle. Dans ce second cas, l'attraction dominait, car l'audiospectateur reconnaissait des titres célèbres et comprenait l'humour du pianiste qui jouait les paroles sous-jacentes aux mélodies contre l'histoire du film. Rick Altman a constaté au début des années 10 une tendance des chansons populaires à fonctionner toujours plus au profit de la narrativisation. Il conclut son chapitre sur la question en disant que «l'utilisation par les pianistes de chansons populaires pour se moquer du film constitue l'un des derniers bastions du cinéma des attractions»<sup>450</sup>. Grâce à la connivence instaurée par le jeu de mots, la médiation entre l'audiospectateur et la représentation est exhibée, alors qu'elle sera occultée (puis supprimée) avec le mode de la (voix)-narration qui se met en place à cette époque.

Pour résumer dans les grandes lignes les bases de l'opposition entre les deux régimes vocaux, je propose de regrouper les aspects traités jusqu'ici au sein de trois grands paradigmes qui, sous-tendus par la dichotomie de l'humain et de la machine, s'attachent à des facettes différentes du déroulement du spectacle.

449. *Ibid.*, p. 136.

450. *Ibid.*, p. 226.

## - MODULARITE VS. FIXATION

La première dichotomie a trait aux implications de la distinction entre la *poésie orale* et l'*écriture* au sens où Paul Zumthor, puis Germain Lacasse conçoivent ces termes. On a pu constater à plusieurs reprises que la généralisation du cinéma parlant avait contribué à accroître l'emprise de l'instance de production sur la distribution des films. À l'époque du «muet», les projectionnistes disposaient par exemple d'une importante latitude d'interprétation en choisissant eux-mêmes la vitesse de défilement du film. Jacques Polet explique à propos du tourneur de manivelle que, «derrière la mécanique du geste, la main du tourneur peut s'abandonner à des *configurations rythmiques* inventives susceptibles de mettre en évidence la contribution performantielle»<sup>451</sup>. Polet décèle dans la participation du projectionniste aux expériences de sonorisation de l'époque «muette» cette même dimension performantielle; il précise à ce propos que «la manivelle instrumentalisait par excellence la fonction essentiellement *attractionnelle* du cinéma des premiers temps»<sup>452</sup>. En soulignant ce terme-clé, Polet nous permet de relever combien les caractéristiques de la voix-attraction s'avèrent communes à d'autres aspects du spectacle cinématographique liés au réglage de la temporalité de la projection. Même si une certaine vitesse de défilement était conseillée et si une forme de standardisation semble s'instaurer vers 1926<sup>453</sup>, les projectionnistes n'en demeuraient pas moins, *en dernière instance* (terme à comprendre tant dans son sens temporel qu'«énonciatif»), seuls maîtres à bord. On peut en dire autant du travail effectué par les bonimenteurs ou les musiciens lorsqu'ils étaient présents dans les salles ou, plus largement, aux décisions de l'exploitant quant à divers aspects de la *présentation* du spectacle. Les modalités du déroulement de la séance constituent un facteur décisif au cours de la première décennie du «cinéma», et surtout durant la période 1895-1902. Dans une synthèse des recherches effectuées par Charles Musser, Tom Gunning rappelle que l'exploitant, en Amérique du moins, prenait à sa charge une grande part des opérations qui relèveront par la suite de la postproduction.

451. Jacques Polet, «Tours et retours de manivelle dans le dispositif cinématographique des premiers temps», in *Le Cinématographe, nouvelle technologie du XX<sup>e</sup> siècle, op. cit.*, p. 151.

452. *Ibid.*, p. 155.

453. Selon Vincent Pinel, «La restauration de films», in *Histoire du cinéma, nouvelles approches, op. cit.*, p. 148.

Ainsi lui incombait-il de constituer un programme, notamment en intercalant des récitations, « en ajoutant de la musique ou d'autres effets sonores et en racontant l'ensemble avec un commentaire parlé »<sup>454</sup>. Étant extérieur à l'objet « film », le son revêt une grande importance dans le contexte de la diffusion : les prologues vivants, la musique, les voix et les bruitages permettent de différencier l'offre, alors que l'image, reproductible, est identique pour toutes les salles. Rick Altman a montré que ce fut avant tout la dimension sonore (chansons illustrées, boniment, effets sonores, etc.) qui représenta pour les exploitants du temps des nickelodéons une variable utile à la personnalisation de leur produit, ce qui ne fut pas sans augmenter la part attractionnelle de la composante sonore : « C'est ainsi que l'exploitation entre 1907 et 1910 tourne de plus en plus au spectacle sonore visant à particulariser l'attraction de chaque exploitant. »<sup>455</sup> Lorsqu'il cite l'exemple du pianiste ivre dont la musique dépend du taux d'alcoolémie de chaque représentation, Kracauer entend mettre en évidence, au-delà de l'anecdote mémorable, cette potentialité d'improvisation et d'influence sur le public que possède tout « interprète » (d'une chanson, d'une musique ou d'un texte écrit)<sup>456</sup>. Étant donné la préséance de l'image dans le spectacle cinématographique, l'univers sonore de la salle constitue le lieu par excellence de la modularité du dispositif.

Les maisons de distribution développèrent d'ailleurs certaines stratégies pour réduire cette part d'aléa. Aussi, elles livrèrent parfois pour la musique d'accompagnement des répertoires (Gaumont édite le *Guide Musical* en 1907) ou, dans les années 20, des « suggestions » musicales qui, adaptées aux différentes séquences des films loués, permettaient de faire « coller » plus précisément la musique au défilement des images<sup>457</sup>. Georges Méliès mettait quant à lui à disposition des bonimenteurs des livrets destinés à faciliter l'élaboration d'un discours verbal relatif à ses films ; selon Gunning, la firme américaine Kalem faisait parvenir des synopsis qui étaient conçus pour

454. Tom Gunning, « Early American Film », in John Hill et Pamela Church Gibson (dir.), *American Cinema and Hollywood, Critical Approaches*, Oxford ; New York, etc. : Oxford University Press, 2000, p. 32.

455. R. Altman, « Naissance de la réception classique... », *art. cit.*, p. 100.

456. S. Kracauer, *Theory of Film...*, *op. cit.*, pp. 137-138.

457. M. Barnier, *En route...*, *op. cit.*, p. 59.

convenir à des présentations orales<sup>458</sup>. De telles démarches ne supprimaient néanmoins ni les contraintes auxquelles les exploitants étaient soumis (par exemple, toutes les salles de province ne disposaient pas de l'orchestre requis par les partitions proposées), ni la marge de manœuvre dont ils bénéficiaient.

Avec la généralisation des techniques du cinéma parlant, ces formes d'interventions extérieures au film qui, dans les années 10 et 20, furent sujettes à des processus d'institutionnalisation – notamment *via* l'influence des rubriques musicales dans les revues éditées par les compagnies de production<sup>459</sup> – disparaissent au profit de leur inscription dans le « texte » filmique même, dès lors passablement plus réfractaire à une éventuelle variabilité. François Albera a souligné le fait que la généralisation du parlant a impliqué l'éviction de l'*oralité*, c'est-à-dire de ce qui appartient à « l'ordre du spectacle, de l'événement, souvent même de l'improvisation », au profit de l'*écrit* au sens large<sup>460</sup>. « L'écrit », c'est autant le texte des intertitres que les voix enregistrées sur la pellicule. Or ce n'est pas un hasard si les cartons de titre et les voix fixées ont favorisé l'institutionnalisation d'une certaine forme de récit que le boniment devait bien souvent mettre à mal en raison de ce que Germain Lacasse considère comme un potentiel de résistance à « l'arraisonnement ».

Sur ce point, un rapprochement semble donc possible entre la voix et cet élément fort discuté à l'époque du cinéma muet qu'est la musique d'accompagnement. À propos des incidences de la présence de musiciens dans la salle sur les modalités du spectacle, il faut remarquer que le point de départ choisi par Noël Burch dans *La Lucarne de l'infini* pour aborder le vococentrisme qui s'est imposé avec l'arrivée du parlant n'est pas celui de la parole, mais de la musique durant la période du « muet ». En effet, il suppose que l'une des fonctions essentielles de l'accompagnement musical a consisté à créer un espace sonore ordonné au sein de la salle de spectacle. En couvrant les bruits parasites, il aurait instauré un rapport plus direct avec le spectateur en tant qu'individu et l'aurait isolé du brouhaha général du public, tant sur un plan acoustique que mental. Cette suppression de perturbations externes aurait

458. Tom Gunning, *D. W. Griffith and the Origine of American Narrative Film: The Early Years at Biograph*, Urbano; Chicago: University of Illinois Press, 1991, p. 93.

459. R. Altman, *Silent Film Sound, op. cit.*, pp. 108-110.

460. F. Albera, « Ouverture », *art. cit.*, pp. 6-7.



permis à chaque spectateur de recentrer son attention sur des pré-occupations « internes », c'est-à-dire relatives à l'univers diégétique. Noël Burch appliquera par la suite cette hypothèse du rôle policier de l'organisation sonore du spectacle à la figure du bonimenteur, considérée comme un « renfort » du « discours musical ». Selon lui, paroles et musique s'inscrivent dans un effort collectif d'embourgeoisement du spectacle cinématographique (selon une évolution qui est celle que prône Stephen Bush dans le passage cité ci-dessus) visant à substituer un « énoncé linéaire contraignant » à « un spectacle *cool*, haché, brouillé, *aléatoire* »<sup>461</sup>. On a vu qu'il s'agissait de relativiser la portée de cette affirmation, Burch sous-estimant l'hétérogénéité des pratiques bonimentorielles. En effet, il analyse le rôle de la parole comme si elle appartenait déjà au MRI, alors qu'elle s'inscrit probablement plus fortement dans l'héritage du spectacle forain, du vaudeville, de la chanson populaire, etc. La fonction du bonimenteur est indissociable des différents contextes dans lesquels il officiait. Par exemple, si G.-M. Coissac témoigne dans son *Manuel du conférencier-projectionniste* de 1908 d'un souci constant pour le maintien de l'ordre dans la salle<sup>462</sup>, c'est parce qu'il envisage les projections comme un instrument pédagogique et non comme un pur divertissement.

Dans *La Lucarne de l'infini*, Burch suggérait que l'utilisation de la musique durant la période muette avait annoncé ce que parachèvera le cinéma parlant, le besoin d'un accompagnement musical relevant d'un « processus global de « naturalisation » du ciné-spectacle, au même titre que le gros plan à venir et, au bout du compte, que la parole synchrone »<sup>463</sup>. C'est la parole synchrone, et non « *over* », que le théoricien considère comme étant le modèle de l'institutionnalisation. Il en va de même chez Jean-Louis Comolli lorsqu'il émet l'hypothèse selon laquelle « la parole et le son synchrones [...] déplacent considérablement le lieu et les moyens jusque-là strictement iconiques de la production de l'impression de réalité », concevant la généralisation du parlant comme l'avènement de la suprématie de l'idéologie « réaliste »<sup>464</sup>. Le texte de Burch et celui de Comolli se rejoignent en raison de bases

461. N. Burch, « Discipline et jouissance », *art. cit.*, p. 89.

462. G.-M. Coissac, *op. cit.*, pp. 176, 184, 188, 204.

463. N. Burch, *La Lucarne...*, *op. cit.*, p. 224.

464. Jean-Louis Comolli, « Caméra, perspective, profondeur de champ, II (5) », *Cahiers du cinéma*, N° 234-235, 1971, p. 98.

marxistes communes qui conduisent ces deux auteurs à déconstruire le naturalisme de «l'art bourgeois». Dans cette perspective, musique et parole participent d'un même phénomène qui implique, pour l'une et l'autre, la soumission du son aux impératifs narratifs et au référent de l'image: la «subordination mécaniste du son à l'image» que Burch décrit (et décrie) relativement à la musique s'imposera définitivement avec l'intégration des voix synchrones. Burch minimise néanmoins la mécanisation croissante des dispositifs en ne distinguant pas l'orchestre présent dans la salle et la musique fixée. En effet, s'il mentionne l'existence d'une tendance inverse à l'institutionnalisation, il ne met pas en évidence la mutation que représente la disparition des musiciens dans la salle au profit de la reproduction mécanique de la musique qui, dans une large mesure, a atténué la composante attractionnelle de l'«accompagnement» pour imposer la «fusion» vers laquelle tendront les dispositifs audiovisuels.

Les premières années de l'interrègne ont toutefois connu certains types d'intervention de la part de l'exploitant que la rapide évolution des techniques et les contraintes économiques ont rendu obsolètes. Ainsi l'abandon progressif du «son-sur-disque», principe de la première version du système Vitaphone, système utilisé par la Warner, au profit d'un «son-sur-film» plus fiable et moins délicat, a-t-il provoqué une diminution des possibilités de modulation du spectacle. En effet, dans le premier cas, le projectionniste tenait encore un peu du «showman» qui doit gérer *live* d'éventuels problèmes de synchronisation<sup>465</sup>. Or cette marge de manœuvre a disparu à la fin de l'année 1930 avec la commercialisation d'appareils permettant de mixer des pistes sonores séparées, et donc de remplacer au stade de la production le travail de l'exploitant qui consistait jusque-là à gérer l'émission des sons en passant du haut-parleur «de fosse», duquel sortait la musique, au haut-parleur situé sur l'écran et réservé aux voix<sup>466</sup>. Avec des systèmes de type Movietone, tout caractère de «performance» disparaît, l'espace de la salle se voyant drastiquement renvoyé à cet oubli nécessaire à l'immersion du spectateur dans l'univers filmique.

On constate que l'opposition entre l'oralité et l'écrit s'applique plus généralement à d'autres composantes que la voix. Ces quelques exemples qui témoignent de l'incidence de la technologie

465. Pour la période 1927-1930 aux États-Unis, voir M. Barnier, *En route..., op. cit.*, p. 44.

466. R. Altman, *Silent Film Sound, op. cit.*, p. 48.

du son sur les conditions de projection des films illustrent l'amoin-  
drissement du pôle attractionnel qu'a connu la représentation  
sonore au sein du dispositif cinématographique. Dans les films qui  
prennent explicitement pour modèles les spectacles de vaudeville,  
de music-hall ou de revue, la modularité n'est pas effective, mais  
représentée, simulée: on a parlé à ce propos d'un régime pseudo-  
attractionnel, défini par une intégration de premier niveau.

L'opposition entre modularité et fixation pourrait laisser penser  
que l'enregistrement phonographique est une condition sine qua  
non de la voix-narration. Toutefois, on a vu que la voix vive du  
bonimenteur pouvait tout aussi bien ressortir à ce pôle. Il s'agit en  
fait de distinguer les modalités de l'émission vocale qui prédispo-  
sent la voix à un certain régime et l'actualisation d'un certain type  
de voix à l'intérieur des possibilités offertes par ces modalités. À cet  
égard, la «fixation» (au sens non plus matériel, mais discursif)  
s'applique au verbal dès lors que celui-ci instaure des liens privilé-  
giés avec la représentation filmique qui, elle, est invariablement  
fixée au niveau matériel. Cette accentuation de la fonction référen-  
tielle s'opère généralement au détriment des fonctions qui sont  
plus directement liées aux instances de la communication ainsi  
qu'au message en tant que tel. En faisant correspondre certaines  
unités linguistiques à des occurrences iconiques qui apparaissent  
invariablement au même endroit et sous la même forme, la parole  
tend à s'inscrire dans le modèle de la «fixation». Naturellement, le  
degré de «fixation» du texte est proportionnel à la subordination  
de ce dernier à l'image. L'émancipation par l'oralité sous la forme  
de l'improvisation, jamais totalement nulle dans le cas de la voix  
vive, conduit à une augmentation de la valeur attractionnelle du  
spectacle. Même lorsque la fonction référentielle est dominante, la  
voix du bonimenteur conserve une part attractionnelle dont le  
public a nécessairement conscience, surtout lorsque survient une  
composante aléatoire, par exemple l'enrouement du locuteur.

Il faut noter que la «fixation» s'applique également à l'organisa-  
tion textuelle de l'énoncé. En effet, le récit obéit à certaines règles  
de structuration sans lesquelles il n'est pas reconnu comme tel par  
le public, alors que le texte proféré par la voix-attraction procède  
plutôt à un éclatement. Dans une structure narrative, les liens  
«chrono-logiques» entre les diverses unités induisent inévitable-  
ment une certaine fixation sans laquelle toute progression serait  
impossible. Comme l'a montré David Bordwell dans *Narration in*

*the Fiction Film*, le spectateur d'un film narratif traditionnel est enjoint à formuler périodiquement des hypothèses concernant la suite de ce qu'il va voir. Comme le spectre couvert par les probabilités des conséquences auxquelles chaque événement représenté se restreint en général avec l'approche de la fin du film, le spectateur oublie les hypothèses qui n'ont pas abouti. Ce travail de sélection de l'information et d'évaluation de la probabilité conduit au cours de la vision du film à « fixer » certaines configurations événementielles ou à retenir des rapports logiques et à en écarter d'autres. L'esthétique des attractions ne réclame par contre pas ce type de réception, mais incite le spectateur à profiter de l'intérêt du moment présent et à minorer l'importance des relations entre les unités. Cette question du *présent* me conduit à clore cette synthèse sur deux autres paramètres, le premier étant lié à l'espace (la *présentation*), le second au temps (la *présentification*).

#### - PRESENTATION VS. DIEGETISATION

On a vu combien tout, dans la voix vive, renvoie à la présence physique de celui qui la profère. De manière récurrente, Paul Zumthor met en évidence les circonstances de la performance vocale, qui sont notamment d'ordre spatial. Sur ce point, la voix-attraction participe d'une organisation audiovisuelle qui relève de la présentation, un mode qui se situe aux antipodes de l'intégration induite par l'usage de la voix-narration. La présentation participe d'une esthétique de l'adresse, tant au niveau de la disposition des éléments profilmiques, du regard et de la gestuelle des acteurs que de la voix (intonation marquée, accentuation exclamative et interrogative, jeu sur la vocalité, etc.) et de la parole (recours aux déictiques). C'est pourquoi la valeur attractionnelle de la voix est d'autant plus forte que le locuteur est visible, ancré dans l'espace nettement délimité du spectacle. Lorsque se manifeste la voix-attraction, l'espace de la salle prime sur l'espace diégétique.

Dans le cinéma parlé, le recul instauré par le bonimenteur renforce le potentiel attractionnel; dans le cinéma parlant, la voix-attraction émane d'un personnage visualisé en voix synchrone. Lorsqu'elle est *over* (ou même « trop » longtemps *off*), la voix se détache de son lieu de profération, substituant à la synchronisation vocolabiale d'autres types de relations, notamment la synchronisation de mise en chaîne qui, grâce à l'ubiquité qu'elle suppose, associe la

voix à un « non lieu » : la voix-narration s'affiche comme totalement solidaire des images, elle ne « présente » pas les images mais donne l'impression de les créer (voir chapitre VI).

Pour Noël Burch, la musique d'accompagnement – que nous avons fréquemment comparée au boniment – constitue dans les années 20 « une survivance têtue de ses origines [celle du cinéma] dans le spectacle populaire et *présentationnel* »<sup>467</sup>. Comme on le verra dans le prochain chapitre, le vaudeville américain et les prologues filmiques en voix synchrone, lieux d'exercice privilégiés de la voix-attraction, peuvent également s'inscrire dans la filiation de ces pratiques présentationnelles.

#### – PRESENTIFICATION VS. REJET DANS LE PASSE

En termes de temporalité, la voix-attraction est ancrée dans un perpétuel présent qui est celui-là même de sa profération. Lorsqu'elle n'est pas reproduite par un enregistrement, la voix confère à chacune de ses actualisations le poids d'un événement singulier qui apparaît au spectateur dans toute son immédiateté. Zumthor envisage ainsi ce phénomène :

« La performance *figure* une expérience, mais en même temps elle *l'est*. Tant qu'elle dure, elle suspend l'action du jugement. Le texte qui se propose, au point de convergence des éléments de ce spectacle vécu, n'appelle pas l'interprétation. La voix qui le prononce ne s'y projette pas (comme le ferait la parole dans l'écriture) : elle est donnée, dans et avec lui, toute-présente. »<sup>468</sup>

La présentification du spectacle induit donc une certaine attitude de la part du locuteur qui s'identifie à l'acte de profération proprement dit. Le terme d'« expérience » permet à Zumthor de faire coexister la représentation et son acte de production. Cette toute-présence ne s'incarne dans le cinéma parlant qu'à travers le renvoi à la manifestation prophonographique de la voix. Le passé auquel la représentation fait référence est donc relatif à la genèse du film, non à l'antériorité des événements évoqués. Le plus souvent, la voix-narration fait par contre de la valeur passée des événements

467. N. Burch, *La Lucarne...*, *op. cit.*, p. 224.

468. P. Zumthor, *Introduction...*, *op. cit.*, p. 234.

auxquels elle se réfère la condition de l'effacement de sa propre inscription dans un présent. Cette question, centrale, de la situation de l'énonciateur (le «narrateur») et de l'usage des temps verbaux nécessite d'aborder plus spécifiquement la voix-narration, comme nous le ferons au chapitre VI, après une étude de cas et des considérations plus historiques sur la notion d'«attraction» rapportée à la composante sonore.

#### IV.8 VOIX-ATTRACTION ET VOIX-NARRATION DANS *LOLA MONTÈS*

*Lola Montès* de Max Ophuls (1955) est un film qui se prête particulièrement bien à l'examen des interactions entre les régimes vocaux puisque sa structure repose à la fois sur une série d'enchaînements – une stratégie de communication narrative qui, au cinéma, conduit le plus souvent à recourir à la voix-over – et sur une référence à un spectacle de type attractionnel. Susan White dit de cette superproduction internationale qu'elle «explore la relation entre cinéma et attraction et entre attraction et narration, dans un contexte précaire entre les pratiques de spectacle traditionnelles et celles d'avant-garde»<sup>469</sup>. Si le terme «avant-garde» ne me paraît pas approprié pour parler du film d'Ophuls, on comprend que Susan White se réfère implicitement au rapprochement proposé par Gunning entre le cinéma des premiers temps et certains films d'avant-garde ultérieurs. L'action de *Lola Montès* se déroule en effet pour partie dans un cirque, c'est-à-dire précisément dans le lieu de spectacle auquel S.M. Eisenstein se référait pour illustrer son concept d'«attractions»<sup>470</sup>.

La première phrase de *Lola Montès*, prononcée par l'écuier (interprété par Peter Ustinov) en introduction au récit de l'ascension et de la chute de la courtisane éponyme, est la suivante: «Et maintenant, Mesdames et Messieurs, l'attraction que vous attendez!» Le mot est lâché, et le mode de l'adresse institué. La formule introductive «et maintenant» suggère que d'autres attractions ont précédé, situées hors du film, et que l'organisation de la représentation obéit au principe de la juxtaposition d'actes autonomes. Lorsqu'il débute son discours, l'écuier prend place devant un vaste rideau comprenant une multitude de petits dessins qui, inscrits

469. Susan White, «*Lola Montès* et le cinéma des attractions», N. Herpe (dir.), 1895, N° 34-35 («Max Ophuls»), 2001, p. 289.

470. S. M. Eisenstein, «Le montage des attractions», *art. cit.*, p. 119.

dans des vignettes, illustrent autant de moments de l'histoire à venir. Ce puzzle visuel rappelle les planches des montreurs d'images itinérants (comme celui que Pabst met en scène au début de *L'Opéra de quat'sous*) qui sont à compter parmi les pionniers des spectacles pré-cinématographiques parlés.

Dans le film d'Ophuls, la voix-attraction intègre des fragments de voix-narration qui lui demeurent subordonnés. L'inverse eût été possible, ainsi que l'a illustré la tentative de Gamma Film, maison productrice de *Lola Montès*. En effet, devant l'échec commercial rencontré par le film, la production entreprit d'exploiter une nouvelle version pour les marchés anglais et américain, intitulée *The Sins of Lola Montès*<sup>471</sup>. La représentation du cirque n'y alternait plus avec les *flash-backs*, mais se voyait rejetée sous une forme écourtée à la fin du film. On ajouta une voix-over dont la fonction consistait à relier les divers fragments audiovisualisés de la vie de la femme fatale. Significativement, cette voix appartient désormais à Martine Carol et remplace celle, tonitruante, du Monsieur Loyal qui domine dans la version originale. À la place de la voix-attraction ostensiblement émise à l'adresse des auditeurs diégétiques et profondément ancrée dans la présence physique du sujet parlant (manifestation phonique du corps et de la corpulence d'Ustinov), les producteurs ont préféré la voix-over de l'héroïne du film, plus « lisse » que celle du personnage de l'écuyer. Cette différence est de taille, puisqu'en raison de son accent anglais mâtiné d'intonation russe, la voix d'Ustinov apparaît dans la matérialité de son grain, et ce d'autant plus que toutes les paroles de l'écuyer ne sont pas dites (ou captées) distinctement lorsque l'acteur se retourne ou s'éloigne de la caméra. La diégèse du film postule en outre que ce personnage s'exprime dans trois langues : une convention veut que cette version destinée au public francophone n'actualise que le discours en français, mais le spectateur est amené à supposer que l'écuyer relate tout ou presque dans les deux autres langues. Il faut mentionner à cet égard que le film d'Ophuls a été conçu un peu comme les « versions multiples » réalisées à l'époque de l'interrègne, puisque de nombreux plans ont été tournés successivement dans trois langues<sup>472</sup>.

471. Voir Alan Larson Williams, *Max Ophuls and the Cinema of Desire. Style and Spectacle in Four Films, 1948-1955*, Salem N.H.: Ayer Co, 1992, p. 141.

472. Pierre Berthomé, « Un caprice qui ne finirait pas... ». La genèse de *Lola Montès*, 1895, N° 34-35, 2001, p. 267.

Au vu des particularités vocales associées à Ustinov, on comprend que le réaménagement conçu pour la version anglophone de *Lola Montès* fasse de la voix-narration le vecteur privilégié de l'intériorisation. Les producteurs semblent avoir imputé la gêne ressentie par le public de la version originale à deux facteurs: d'une part au caractère discontinu du film, d'autre part au décentrage de l'origine énonciative *via* l'intervention du «bonisseur» de cirque. Familier des pratiques de distanciation qu'il ne cessera d'expérimenter sur scène<sup>473</sup>, Max Ophuls avait prévu de faire jouer à l'écuyer le rôle d'un narrateur «juxtadiégétique», c'est-à-dire d'une instance énonciative qui se situe dans un rapport d'extériorité radicale à la représentation visuelle. Cette extériorité n'est pas sans rappeler la relation qu'entretenait le bonimenteur du «cinéma» parlé avec les images projetées. On comprend dès lors que la matérialité de la voix soit exhibée avec autant de force dans *Lola Montès*: ce film est l'une des rares réalisations de la période dite «classique» qui, tout en n'appartenant pas au genre de la comédie musicale, privilégie explicitement la voix-attraction et permet simultanément un embryon de voix-narration. Il n'est pas étonnant qu'un cinéaste comme François Truffaut, fervent défenseur de l'œuvre cinématographique de Guitry (voir chapitre III), fit l'éloge du film d'Ophuls dans un article des *Cahiers du cinéma* en grande partie consacré à la question de la voix, où il note le «double et le triple *décalage* qu'il y a constamment dans *Lola* entre la personnalité des acteurs et leur diction, entre leur diction et leur texte»<sup>474</sup>. Le film d'Ophuls appartient à cette catégorie de films qui travaillent à même la matière vocale en jouant de la tension qui s'établit entre la voix-attraction et la voix-narration.

Les commentaires de l'écuyer ne sont pas adressés au spectateur du film, mais à son «délégué» diégétique qui, comme lui, demeure constamment dans l'ombre. Précisons que la caméra ne se trouve pas pour autant à l'emplacement présumé du public du cirque: comme à son habitude, Ophuls effectue de nombreux mouvements d'appareil (dont certains sont marqués sur le plan de la monstration) et varie sans cesse la distance qui nous sépare de l'objet principal du regard, ce corps féminin exhibé en «bête de foire» et asservi à la voix masculine d'un écuyer muni d'un fouet.

473. Voir Helmut G. Asper, «De la fosse du souffleur au micro de l'écran», *1895*, N° 34-35, 2001, p. 56.

474. François Truffaut, «Lola au bûcher», *Cahiers du Cinéma*, N° 55, 1956, p. 30.



Nous sommes donc aux antipodes d'une construction frontale de l'espace qui renverrait, comme dans certains prologues évoqués précédemment, à une forme conventionnelle de théâtralité. Reprenant l'idée du carrousel de *La Ronde* (1950), Ophuls utilise l'espace circulaire de l'arène pour induire une multiplicité de « points de vue » et leur perpétuelle transformation. La chute finale de Lola est d'ailleurs filmée en plan subjectif, ce qui soustrait au regard du spectateur du film une action qui constitue, pour le public diégétique, le clou du spectacle. À mon sens, cette occultation résulte d'une narrativisation qui fonctionne exclusivement à l'intention du public du film : sachant qu'un médecin a demandé au propriétaire du cirque d'installer les filets de sécurité en raison du mauvais état de santé de Lola, le spectateur de *Lola Montès* a accès, par le truchement du montage alterné, à des éléments d'intrigue qui portent sur la vie privée actuelle de la femme à scandale. Tout ce qui a trait à cette dimension-là du récit est transmis sur le mode traditionnel de la voix-action, et cela même dans la bouche du Monsieur Loyal qui, bien que recourant majoritairement à la voix-attraction, intercale dans ses harangues clamées à la foule des phrases qu'il chuchote à Lola. Les insertions répétées de *flash-backs* et les nombreux mouvements d'appareil complexes perturbent l'appréhension de l'espace-temps scénique. À cet égard, Jacobsen dit de *Lola Montès* qu'il « est structuré comme un show télévisuel moderne avec présence du public dans le studio »<sup>475</sup>. Cette comparaison pointe l'originalité du film qui est de créer, grâce à la voix de l'écuyer, une pseudo-interactivité avec le public tout en exhibant fortement l'artifice de la représentation scénique, puisque le spectateur du film dispose d'un pouvoir ubiquitaire qui lui permet d'accéder aux loges et aux coulisses où s'activent les machinistes. Quant aux spectateurs présents dans le film, ils existent prioritairement par la voix, mais celle-ci est systématiquement *off*. Ophuls déclare s'être inspiré de certains programmes radio-phoniques pour rédiger le texte des demandes de ce public avide de révélations sur la vie privée de Lola<sup>476</sup>. La référence à ce médium semble avoir influencé la conciliation opérée dans *Lola Montès* entre le statut *off* ou *over* de la voix et sa composante attractionnelle.

475. Helmut G. Asper et al. (dir.), *Max Ophuls*, Munich: Hanser Verlag, 1989, p. 242.

476. Cité par A. L. Williams, *art. cit.*, p. 140.

Le caractère attractionnel des voix de *Lola Montès* se manifeste dans leur matérialité même: Ustinov entonne la chanson des «douze perfections», et ses paroles sont parfois reprises sur un mode grotesque et d'une voix nasillarde par un nain. De plus, les voix des figurants du cirque sont pour la plupart altérées, renvoyant ainsi à l'aspect factice de cette grande mascarade. Tel est par exemple le cas du groom qui, le visage dissimulé derrière un masque, répond au médecin d'une voix étouffée, rauque et gutturale, alors que sa stature laisse penser qu'il s'agit d'un enfant. Quant à la parole de Lola, elle est réduite à de brèves «citations» dans un texte proféré par Ustinov, à l'instar des dialogues d'un roman qui, provisoirement, interrompraient la «voix» d'un narrateur. Non seulement l'écuyer parle de Lola, mais il parle *pour* elle, à l'instar de certains bonimenteurs du cinéma des premiers temps qui doubleraient les personnages de l'écran.

La parole de l'écuyer constitue le vecteur principal de la transition vers les enchâssements. On peut donc considérer que ce personnage use par intermittence d'une voix-narration inscrite dans le régime vocal dominant de la voix-attraction. Sa position intermédiaire assure à la voix une double fonction de médiation: entre la représentation et le public d'une part, entre différents modes de représentation d'autre part. Le personnage qu'interprète Ustinov ne se contente pas de raconter l'histoire de Lola, mais tente de la rendre captivante, par exemple en exacerbant l'«authenticité» des actions relatées, ou en adoptant parfois un ton grandiloquent. La parole contribue à narrativiser *Lola Montès* de deux manières: d'une part, elle rend compréhensible, dans la situation-cadre, une série d'actions représentées sur scène qui paraissent passablement chaotiques et demeurent muettes; d'autre part, elle permet d'introduire certains enchâssements. Dans les deux cas, le personnage interprété par Peter Ustinov utilise le langage verbal pour linéariser une consécution d'événements discontinus au niveau de la bande-image. En effet, en raison de la nécessaire schématisation qu'impliquent l'expression pantomimique et les décors du cirque, plusieurs numéros de l'arène construisent leur référent de façon trop partielle pour être compris par le spectateur sans l'apport du commentaire verbal de l'écuyer. En outre, le filmage des performances des acrobates ne favorise pas la lisibilité des actions, soit parce que la distance spatiale entre ceux-ci et la caméra est trop grande, soit parce que des éléments profilmiques ou des caches

obstruent la vision. On peut affirmer qu'en l'absence de sons, les numéros perdent toute dimension narrative.

Or la narration de ce Monsieur Loyal est fortement empreinte d'oralité, ainsi qu'en témoignent la récurrence des adresses et la forme du « résumé ». Le sociolinguiste William Labov tient cette dernière pour une caractéristique essentielle du récit oral intégré à une conversation ordinaire : il s'agit d'un ensemble de propositions qui, par avance, indiquent le contenu principal de l'histoire<sup>477</sup>. Cette anticipation qui pourrait paraître tout à fait incongrue dans un récit littéraire traditionnel, où il s'agit au contraire de ménager des surprises en égrenant progressivement des rebondissements nodaux inscrits dans une courbe évolutive, s'explique dans le cas de la situation conversationnelle par la nécessité que rencontre le narrateur de motiver la rupture qu'il introduit dans le dialogue en insérant son histoire, c'est-à-dire en embrayant sur une communication de type monologique. Cette stratégie discursive est bien sûr simulée dans le cas de l'écuyer de *Lola Montès*, mais elle permet à mon sens de rendre compte de l'interruption provisoire du régime dominant de la voix-attraction. En effet, le « résumé » est un moyen de souligner la valeur attractionnelle des événements que le locuteur s'apprête à conter et, simultanément, d'indiquer le passage à l'énonciation narrative. Dans les répliques de l'écuyer, cette fonction de résumé se combine fréquemment avec celle du *titrage* des parties qui convoque un référent intertextuel particulier tout à fait adapté au mélange d'attraction et de narration dont le film procède : les manchettes des journaux à scandale. Enclin à entrer dans les détails scabreux de faits-divers, l'écuyer réduit dans un premier temps le contenu narratif à une structure minimale qui, formulée avec concision, convient à la nature attractionnelle de la représentation. Sarah Kozloff souligne dans son ouvrage sur la voix-over la pertinence du modèle de Labov en notant qu'au cinéma nous sommes en présence de récits oraux qui, du point de vue strictement verbal, sont généralement très fragmentaires<sup>478</sup>. La situation dialogique traitée par Labov me paraît toutefois fort différente de la voix-narration du cinéma que j'aurais plutôt tendance à rapprocher à cet égard (surtout sous la forme de la voix-over) du récit scriptural. Ainsi la phase de « résumé » me semble concerner

477. William Labov, *Le Parler ordinaire. La Langue dans les ghettos noirs des États-Unis*, Paris : Éditions de Minuit, 1978, pp. 299-300.

478. S. Kozloff, *Invisible Storytellers*, op. cit., pp. 3-4.

plus spécifiquement les énoncés relevant de la voix-attraction ou de la voix-action, ou constituer un indice de celles-ci au sein du régime de la voix-narration.

Ces remarques sur la valeur attractionnelle de la voix de l'écuyer ne doit pas faire oublier qu'elle est par ailleurs l'instigatrice des transitions vers le passé. On verra, lorsque nous aborderons la question des prologues, combien la parole est prédisposée à endosser au cinéma une fonction d'instauration d'un monde (audio)visualisé. Dans *Lola Montès*, les mots constituent un facteur décisif pour nous emporter vers un autre lieu et un autre temps. Ce n'est pas sans raison si plusieurs recenseurs du film l'ont spontanément rapproché, à sa sortie, de *Citizen Kane* et de *La Splendeur des Amberson*<sup>479</sup> : le film d'Ophüls fait resurgir un passé, et cette réminiscence est partiellement médiatisée par une voix.

Même si de nombreux moments-clés de la vie de la courtisane sont élidés, les *flash-backs* présentent un degré de narrativité supérieur à la plupart des épisodes représentés dans l'arène. Il est nécessaire de s'interroger sur l'intervention de la voix par rapport à leur insertion. En effet, comme la narration verbale scénique est presque entièrement placée sous la responsabilité énonciative de l'écuyer, on peut se demander si cette position de contrôle discursif se perpétue lors des séquences audiovisualisées. Kaja Silverman remarque à ce propos : « Parce que sa voix [celle de Peter Ustinov] parle de temps à autre par-dessus les images de Lola, il semble moins ancré diégétiquement, et donc plus proche du point de l'origine textuelle. »<sup>480</sup> L'explication de la théoricienne s'inscrit dans une certaine conception anthropomorphisante de l'instance de l'énonciation filmique que je discuterai en fin d'ouvrage. On peut relever pour l'instant que Silverman associe la place occupée par l'écuyer à celle que l'on attribue généralement aux voix dont « l'ancrage diégétique » est atténué par l'absence de synchronisation vocolabiale. Tout d'abord, il faut souligner qu'Ustinov est fréquemment hors-champ dans les plans du cirque. Le statut *off* de sa voix le prédispose donc effectivement à intervenir en voix-over, c'est-à-dire avec un degré supérieur d'extériorité par rapport à l'image. Toutefois, on constate que sur les sept transitions vers le

479. Voir les *Cahiers du Cinéma*, « Le dossier de presse de *Lola Montès* », N° 55, janvier 1956, p. 56.

480. Kaja Silverman, *The Acoustic Mirror. The Female Voice in Psychoanalysis and Cinema*, Bloomington ; Indianapolis : Indiana University Press, 1988, p. 57.

passé que compte *Lola Montès* (dans sa version française actuellement distribuée, d'une durée d'une heure cinquante), seules deux sont véritablement portées par le personnage d'Ustinov, dans le sens où sa voix-over enjambe les deux mondes. Il s'agit tout d'abord de l'introduction au séjour de Lola sur la Côte d'Azur, dont la responsabilité énonciative semble motivée en termes de focalisation puisque le Monsieur Loyal intervient lui-même dans l'épisode conté, et cela dès la première image de l'enchâssement; le second cas est celui de l'épisode bavarois, de loin le plus long des *flash-backs* puisqu'il dure plus d'un tiers de la durée totale du film. Le lieu des interventions vocales de l'écuyer par rapport à la chronologie des enchâssements laisse entrevoir le mouvement du film qui connaît une emprise croissante de cette instance verbale diégétique sur la façon dont le passé est évoqué.

Ophuls renforce l'opposition entre représentation scénique (attractions) et audiovisualisations filmiques (narration) par un emploi contrasté de l'Eastmancolor: ce procédé est utilisé de manière beaucoup plus sobre dans les enchâssements (où l'on repère tout de suite une teinte dominante) que dans la situation-cadre, qui comprend des couleurs extrêmement vives et bigarrées. Un plan ambigu en termes de niveau narratif – on passe d'une image du ciel nocturne que contemple Lola depuis le pont d'un bateau à des étoiles grossièrement dessinées sur une toile, puis cachées par un rideau qui se ferme sur la scène – montre que l'émergence de la voix-over emporte avec elle quelque chose du monde du cirque auquel, en tant que voix-attraction (synchrone), elle est initialement attachée. Le spectateur est amené à accepter, par convention mais aussi parce que certains éléments sont communs à ces deux « mondes » (à l'exemple du carrosse de l'épisode de Liszt qui, après avoir joué un rôle central dans l'enchâssement, apparaît dans l'arène), que l'écuyer continue à bonimenter le spectacle durant l'audiovisualisation en *flash-back*. Ainsi que l'a noté Kaja Silverman, le passé de Lola devient alors une sorte de prolongement de la parole de l'écuyer.

Néanmoins, d'autres voix féminines se font entendre pour motiver l'insertion de *flash-backs*. Le premier enchâssement s'effectue par la répétition échoïque et irréaliste d'une partie de la question d'une spectatrice qui, s'exprimant *off* (sur l'injonction de l'écuyer), ne donne lieu à aucun ancrage. L'interrogation « s'en souvient-elle? » est réitérée six fois alors qu'un lent fondu enchaîné

occasionne la surimpression progressive des images du passé sur un gros plan de Lola. Martine Carol se tient au centre de l'image, l'air pensif, de sorte que l'occurrence vocale s'interprète comme un instrument d'introspection. Le spectateur comprendra ensuite, avec le dévoilement d'un décor correspondant à l'action du *flash-back*, qu'il ne s'agissait pas seulement du «souvenir» de l'héroïne déchuée, mais de la représentation cinématographique d'une performance de cirque. À cette intériorisation fait suite une phase intermédiaire où le personnage interprété par Peter Ustinov n'a pas encore totalement pris fictivement les rênes du récit, car c'est Lola qui relate verbalement son passé. Toutefois, la jeune femme apparaît totalement extérieure à sa propre histoire dont elle a été dépossédée par le *show business*. Elle n'a pas d'emprise sur ses paroles: elle bafouille, laisse l'écuyer lui suggérer la suite de son texte ou parler à sa place. Alors qu'au début du film, la voix d'une spectatrice indiscreète était utilisée pour atteindre les pensées de Lola, c'est désormais la voix de celle-ci qui se soumet à la représentation d'un autre. Le vécu devient récit. Lorsqu'elle tente de reprendre le fil de l'histoire après le *flash-back*, l'écuyer l'interrompt et prend en charge l'énonciation narrative (fictive) jusqu'à la fin du film, contrebalançant la part attractionnelle de ses interventions dans un savant équilibre qui caractérise l'hybridité du régime vocal de *Lola Montès*.

## V. LA VOIX-ATTRACTION DANS L'HISTOIRE DU CINEMA

### V.1 LES SONS COMME «ATTRACTIONS» DANS LE CINEMA DES PREMIERS TEMPS

La notion d'«attraction» implique selon moi trois types de considération interdépendants que je dissocierais pour des raisons de clarté méthodologique: elle a premièrement trait à un niveau *technique*, puisque toute innovation se voit passagèrement considérée par le public comme une «attraction» au sens le plus courant du terme (les spectateurs sont avant tout captivés par l'appareillage, les films ne servant qu'à exhiber son fonctionnement); deuxièmement, sur un plan *esthétique*, le «cinéma des attractions» témoigne d'un mode de représentation qui se distingue des pratiques ultérieures ou dominantes; enfin, à un dernier niveau étroitement lié au précédent, le «cinéma des attractions» désigne un ensemble de conditions caractérisant la *réception* spectatorielle, et donc postule la mise en place d'un certain dispositif. On retrouve là des paramètres que j'ai abordés à propos du bonimenteur du cinéma des premiers temps, puisque l'instance humaine peut, dans le dispositif «cinéma», se mettre au service du régime des attractions. Les directeurs du numéro 22 de la revue *Iris* consacré au bonimenteur terminent d'ailleurs leur éditorial en suggérant que «la véritable attraction du *cinéma des attractions* n'est pas toujours dans l'image, elle est parfois devant elle ou à ses côtés»<sup>481</sup>. Cette formulation renvoie autant aux théories récentes qu'à la signification que l'on pouvait donner au terme «attraction» à l'époque. Ainsi Le Fraper disait-il par exemple en 1911 du bonimenteur: «S'il est instruit, intelligent et travailleur, il ne tardera pas à devenir une attraction.»<sup>482</sup> Après

481. André Gaudreault et Germain Lacasse, *Iris*, N° 22, 1996, p. 8.

482. Charles Le Fraper, «Les Conférenciers», *Le Courrier cinématographique*, 23 septembre 1911, cité en annexe de A. Gaudreault, «Le retour du [bonimenteur] refoulé...», *art. cit.*, p. 29.

avoir discuté au chapitre précédent l'acception théorique de ce terme, il me paraît intéressant de s'interroger sur son usage au début du XX<sup>e</sup> siècle et dans la phase initiale de la généralisation du parlant.

Durant les premières années de l'exploitation cinématographique, on utilisait déjà les sons pour conférer à la séance une valeur attractionnelle. Un programme de 1899 présentant les «Attractions des Grands Magasins Dufayel» mentionne par exemple le «Cinématographe Lumière, avec imitations parfaites des bruits»<sup>483</sup>. Parmi les exemples cités, on trouve, en caractères plus épais, la mention «Scènes animées et parlées». Si les bruitages étaient sans doute réalisés en coulisses et s'inscrivaient donc dans une tradition théâtrale, l'émission vocale devait bénéficier d'une production mécanisée grâce à l'apport d'un phonographe. En effet, l'utilisation de cet appareil s'avère d'autant plus probable qu'une attraction mentionnée dans la suite du programme consistait justement en une présentation du Stentor, un type de gramophone. Il est intéressant de noter que cet appareil est annoncé dans cette publicité comme une attraction indépendante et qu'il bénéficie, en tant que «nouvel appareil créé par la Société Pathé frères», d'une importance graphique (grandeur et épaisseur des caractères, centralité de la mention) égale à celle du cinématographe. Côtéant sur la même affiche d'autres expériences menées grâce à des technologies nouvelles, phonographe et cinématographe font donc office, pour leur promoteur, d'attractions scientifiques. C'est dans ce sens qu'il faut comprendre le couplage de l'un et de l'autre: l'effet illusionniste éventuellement induit par le synchronisme dont se souvient Henri Diamant-Berger qui, enfant, avait assisté à ces projections sonores chez Dufayel («un phonographe – mais il me semblait que c'était le personnage lui-même – exhalait sa douleur en chantant»), devait probablement céder le pas à l'admiration face à la perfection technique de l'attraction<sup>484</sup>. Si, dans le catalogue de la firme Pathé de 1904, ces «scènes ciné-phonographiques» sont classées parmi les «attractions merveilleuses», ce n'est pas en vertu du contenu de ce qui est montré (comme on le dirait de féeries), mais du caractère exceptionnel de la technique utilisée<sup>485</sup>.

483. J.-J. Meusy, *op. cit.*, p. 37.

484. *Ibid.*, p. 102.

485. Édouard Arnoldy, *Une histoire du cinéma parlant et sa diagonale mobile. De l'art cinématographique, des «films chantants et parlants» et de leurs séries*, tome 2, thèse de doctorat, Université de Liège, 2001, p. 127, note 2.



La nature du représenté peut toutefois fournir un potentiel attractionnel particulièrement propice à une exploitation par le cinéma sonore. Francesco Bono a relevé que les *Biophon Tonbilder*, bandes synchronisées avec un gramophone (réalisées en Allemagne à partir de 1907 par Oskar Messter), consistaient parfois en des extraits d'opérettes filmées, le pionnier allemand usant de la « combinaison de l'image, du chant et de la musique pour impressionner son public »<sup>486</sup>. L'historien remarque en outre que, lorsqu'un acteur comme Alexander Girardi apparaît dans ces bandes sonorisées, c'est en tant que star d'opérette, non en tant que personnage, comme en témoignent deux titres de 1908 qui comportent son véritable nom<sup>487</sup>. Le film fonctionne donc comme rappel de la scène, et ne construit pas d'univers diégétique clos. La voix-attraction exhibe une performance qui est à la fois celle du chanteur et celle de la technique utilisée pour la reproduire. Les spectateurs sont fascinés par l'intégration de l'humain dans la machine et, inversement, par la capacité de l'innovation technique à produire une représentation de l'humain. Le système des acteurs dissimulés derrière l'écran ne présente par exemple rien de nouveau pour le public du début du siècle, mais constitue tout de même une attraction dans la mesure où la nouveauté naît de la conjonction de locuteurs et d'un mécanisme de projection auquel ceux-là doivent se conformer à des fins de synchronisation.

## V.2 L'ATTRACTION SONORE A LA FIN DES ANNEES 20 : LE CONSTAT D'ALEXANDRE ARNOUX

Pour les détracteurs du parlant de la fin des années 20, la technologie sonore menaçait le cinéma d'une régression au « théâtre filmé », alors que le cinéma muet avait atteint selon eux son « apogée ». Cette crainte récurrente face à un appauvrissement des moyens d'expression cinématographique qui résulterait des contraintes de la prise de son ne se vérifie cependant pas dans la production du début des années 30 où, comme l'a notamment montré Martin Barnier, le son fut rapidement exploité dans ses potentialités propres. En 1985, sous l'impulsion des travaux de

486. Francesco Bono, « Glücklich ist, wer vergisst... Operette und Film. Analyse einer Beziehung », in Katja Uhlenbrok (dir.), *Musik, Spektakel, Film, Musiktheater und Tanzkultur im deutschen Film 1922-1937*, Munich: Editions text+kritik, 1998, p. 31.

487. *Ibid.*, p. 44, note 4.

Rick Altman, William Johnson s'opposait déjà au lieu commun de la régression en constatant que les réalisateurs, en 1929 déjà, «explorèrent de nombreuses façons de combiner le son et l'image» et «s'adaptèrent rapidement à la bande-son»<sup>488</sup>. Toutefois, les caractéristiques communément associées au «théâtre filmé» (frontalité, fixité du cadre, quasi-absence de montage, ...) valent effectivement pour des réalisations parlantes antérieures à la généralisation du son, comme dans les Phonoscènes Gaumont ou dans les premières bandes Vitaphone. La technique, qui sera certes perfectionnée (au niveau des microphones, de l'insonorisation de l'appareil de prise de vues, etc.), n'est pas seule en cause, car elle fut, de toute manière, mise au service d'une esthétique héritée des arts de la scène comme le vaudeville, le music-hall et l'opéra. Les pionniers du son à l'époque du «muet» projetaient avant tout de faire entendre et voir à un public considérablement élargi des chanteurs célèbres. C'est d'ailleurs sur cette démocratisation de la culture musicale – qui connaît également un avantage non négligeable dans la vente de disques – que Will Hays met l'accent dans son discours d'introduction aux *Vitaphone Shorts*.

Dans les bandes du début du parlant, pour une bonne part des «chansons illustrées», l'absence de travail au niveau filmographique et une disposition du profilmique imitant souvent la scène théâtrale apparentent les «attractions» audiovisuelles à d'autres types de spectacles. Leur intérêt ne réside pas tant dans ce qu'elles montrent que dans le fait même qu'elles présentent du son, c'est-à-dire dans ce qui en fait une innovation. Dans les premiers temps de la généralisation du cinéma sonore et parlant, l'invention technique possède à elle seule une valeur attractionnelle. Ainsi Michel Collinet intitule-t-il, en mars 1928, un article sur le cinéma parlant «Les nouvelles possibilités d'attraction cinématographique». Il envisage explicitement les potentialités de ce qu'il nomme le cinéma «acoustique» au sein du paradigme de l'«attraction», dans l'acception première du terme qui, selon Jean Giraud, renvoyait à un «spectacle curieux», à l'«élément captivant, sensationnel d'un programme»<sup>489</sup> :

488. William Johnson, «The Liberation of Echo: A New Hearing for Film Sound», *Film Quarterly*, Vol. 38, N° 4, 1985, pp. 2-3.

489. Jean Giraud, *Le Lexique du cinéma français: des origines à 1930*, Paris: CNRS, 1958, p. 48.

«Nous avons vu que le cinéma a dévoré en peu d'années les sujets accumulés par des siècles de littérature: qu'il a usé et abusé des clous attractifs comme éruption de volcan, naufrages, catastrophes de tous genres, courses, bals, fêtes, etc. Et qu'il fallait craindre une désaffection du public lassé. Nous cherchions *cette autre chose* [...]. Cette *autre chose*, cette attraction qui renouvellera le spectacle cinématographique, qui obligera tous les curieux à venir et revenir dans les salles, le tenons-nous? Il semble que les efforts de la maison Gaumont nous permettent bien de répondre oui.»<sup>490</sup>

Si, pour Collinet, le parlant importe pour l'avenir du cinéma, c'est avant tout parce qu'il permettra de gagner un nouveau public, de susciter la curiosité grâce à l'innovation technique qu'il propose. L'auteur de l'article ne conçoit pas la possibilité de longs métrages parlants – précisons qu'il n'a pas encore pu voir *Le Chanteur de jazz*, qui ne sortira à Paris que dix mois plus tard – mais uniquement de brefs intermèdes, des «clous» qui rappellent tout à fait le mode de présentation des films du début du siècle. En effet, Collinet prend en compte le programme dans son intégralité et envisage positivement l'insertion de courtes bandes parlantes qui «pourront fournir au spectacle cinématographique un élément nouveau d'attraction» et grâce auxquelles «l'ensemble de la soirée deviendra plus gai, plus vivant, plus aimable»<sup>491</sup>. Ce sont bien les spectacles «vivants» que le journaliste semble évoquer (par exemple les sketches et *playlets* du vaudeville) lorsqu'il imagine le futur rôle du parlant. Enthousiaste face à ce renouvellement, il montre que la curiosité pour la technique peut primer sur les considérations esthétiques.

Comme il n'est pas possible d'évoquer ici toutes les figures majeures de la réception critique de cette époque, je propose d'étudier les recensions d'un commentateur de premier plan, dont les remarques attestent du lien que je suppose entre les premiers parlants et le cinéma des premiers temps. Il s'agit d'Alexandre Arnoux, rédacteur en chef de l'hebdomadaire de cinéma *Pour Vous*, qui fut l'un des principaux représentants de la réception française du cinéma parlant au sein de la presse spécialisée<sup>492</sup>.

490. Michel Collinet, «Les nouvelles possibilités d'attraction cinématographique. Le cinéma parlant», *Le Cinéopse*, N° 103, 1<sup>er</sup> mars 1928, p. 269.

491. *Ibid.*, p. 270.

492. Voir Philippe d'Hugues, «Au commencement était Adam», in Noël Herpe et Emmanuelle Toulet (dir.), *René Clair ou le cinéma à la lettre*, Paris: AFRHC, 2000, p. 40.

Dans un bref historique rédigé en 1934, Arnoux propose une périodisation basée sur le dépassement d'un premier stade qu'aurait constitué la mise en spectacle de l'invention technologique elle-même. Il qualifie cette mutation de « passage de l'attraction lancée à grand renfort de publicité à un degré supérieur, à une sphère plus haute où l'art déjà dispute la suprématie au tour de main »<sup>493</sup>. Dans cette affirmation émise à la fin de l'interrègne – c'est-à-dire, dans l'évolution de la réflexion d'Arnoux, quelques années après son refus du cinéma parlant, justement lié en 1928 à ce qu'il considère ici comme la « basse » sphère des premières productions parlantes ou sonorisées, l'auteur pose une distinction entre les premières réalisations qualifiées d'« attractions » et les productions « artistiques » plus tardives. Cette dévalorisation de la composante attractionnelle, comprise dans son sens tout à fait littéral (c'est l'invention en elle-même qui « fait attraction »), n'est pas sans parenté avec la relative éviction des spécificités du cinéma des premiers temps que l'on constate chez les historiens traditionnels du cinéma. Arnoux utilise d'ailleurs l'expression « confusion de foire » pour dépeindre le climat de cette transition située entre 1927 et 1929, une formule qui, dans le contexte actuel des recherches en histoire du cinéma, évoque également les circonstances de l'exploitation des premières années du cinématographe. La connotation péjorative dont Arnoux affuble le terme « attraction » participe d'un jugement esthétique plus global qui place les premiers films parlants aux antipodes de l'« art ». Si Arnoux fait par la suite mention de la disparition de l'attraction au profit d'un « souci *du style* », c'est bien qu'il institue en norme *un style* en particulier. Et la narration, comprise au sein du couple attraction/narration comme une esthétique de la continuité et de la transparence, aura bien vite (re)pris le pas sur les attractions. Car ce qui déçoit Arnoux, dans les premières bandes parlantes qu'il a l'occasion de voir – du moins tel qu'il en rend compte dans son texte postérieur destiné à livrer les réactions d'un des premiers témoins qu'il fut – c'est surtout l'exhibition même du dispositif à travers ses imperfections (« émission assez ridicule de la voix qui paraissait provenir d'on ne sait quel arrière-plan rauque ») ou des choix de mise en scène (« articulation détaillée des comédiens dont les

493. Alexandre Arnoux, *Du muet au parlant. Mémoires d'un témoin*, Paris: La Nouvelle Édition, 1946 [1<sup>re</sup> édition 1934], p. 84.

muscles de la face, trop rapprochés de l'objectif, jouaient une pantomime désespérée et fort pénible)»<sup>494</sup>.

De telles remarques de la part d'un contemporain de l'invention nous permettent de relativiser le caractère évident d'un supposé renforcement de l'impression de réalité *via* la synchronisation entre la mimique articuloire visualisée et la voix, même s'il s'agissait là de l'objectif de tous les concepteurs de «machines parlantes». Tant que l'utilisation du son participe d'une logique attractionnelle, que cela soit par respect d'une certaine tradition esthétique ou par nécessité technique, la représentation audiovisuelle résiste à ce que Gunning nomme «l'absorption diégétique» du spectateur. Le terme «exhibition» apparaissait d'ailleurs déjà, fin 1929, sous la plume d'Alexandre Arnoux: «Il [le public français] avait vu, je sais bien, *Le Chanteur de jazz* et *Les Innocents de Paris*, mais ces bandes constituaient plutôt des *exhibitions* de vedettes chantantes que des *talkies*, plutôt des enregistrements phonographiques agrémentés d'images que des actions visuelles.»<sup>495</sup> En poursuivant par l'affirmation qu'il «n'en va plus de même aujourd'hui», Arnoux développe la même argumentation que dans sa conférence de 1934 en distinguant une première phase d'expérimentations et l'avènement «d'œuvres d'art». Ce qu'il énonce à propos des «vedettes chantantes» serait d'ailleurs également valable pour les *Vitaphone Shorts* ou pour des bandes antérieures telles que les Phonoscènes, où les stars sont «exhibées» sans donner corps à un personnage agissant. L'opinion d'Arnoux se base sur une comparaison avec les productions les plus récentes, dont l'aspect attractionnel tend à s'amenuiser avec l'institutionnalisation qui est en cours. Ainsi, ce qu'il dit du *Chanteur de jazz* ne correspond pas à ce qu'il écrivait au début de la même année, juste après la première présentation du film en France: «La ligne générale, cependant, demeure fort nette; le développement se poursuit avec une logique évidente; jamais ou presque jamais nous avons eu l'impression, ce que je craignais d'abord, d'une mosaïque de clous et d'*attractions* techniques.»<sup>496</sup> Les deux dernières expressions renvoient indubitablement à certaines formes de spectacles vivants populaires dont les productions cinématographiques sont censées, selon Arnoux, se distinguer.

494. *Ibid.*, p. 86.

495. *Pour Vous*, N° 51, 7 novembre 1929, cité par R. Icart, *La Révolution du parlant*, op. cit., p. 199.

496. *Ibid.*, p. 221 (*Pour Vous*, N° 11, 31 janvier 1929).

Contre l'aspect fragmentaire et hétérogène de la « mosaïque », le critique valorise la « logique » d'un « développement », c'est-à-dire l'unicité d'un récit.

On voit combien l'opposition entre le cinéma des attractions et l'intégration narrative discutée par Gaudreault et Gunning constitue un enjeu déterminant parmi ceux qui pensent et évaluent l'innovation technique aux débuts de la généralisation du parlant. Le lien postulé par Arnoux entre le statut d'attraction et les « balbutiements » provisoires d'une nouveauté technologique appartient par contre aux lieux communs d'une conception évolutionniste de l'histoire du cinéma, qui voit dans la « spécificité » du médium, et donc dans son affranchissement de pratiques hybrides, la condition de possibilité de son statut artistique.

Non seulement Arnoux représente une figure importante du champ cinématographique de la France de l'époque – il est notamment scénariste de *L'Atlantide* (Georg Wilhelm Pabst, 1932), film réalisé en multiversions (française, allemande et anglaise) qui présente un récit enchâssé initié par une voix d'abord acousmatique, puis momentanément *over*, mais son avis n'est en rien marginal à la fin des années 20. Par exemple, le critique contemporain Jean Tedesco motive son opposition à un certain type de cinéma sonore – dans lequel on pourrait dire que prédomine la voix-attraction – par des arguments semblables. Ainsi prophétise-t-il l'évolution suivante : « Ces expériences de film parlé se résumeront à des morceaux, le plus souvent chantés, insérés dans les films, ou bien à des espèces d'*attractions*, soi-disant plus vivantes, où l'on retrouvera le mauvais goût des présentations américaines des grandes salles de Broadway ou de Paris. »<sup>497</sup> On perçoit chez Tedesco le même jugement relatif aux spectacles populaires opposés à l'art, la même crainte devant une utilisation de la voix et du chant comme facteur de destruction de l'unité du spectacle, alors qu'à partir de 1915 environ cette unité s'est vue consolidée par l'adoption généralisée du format long métrage. Il est à cet égard intéressant de relever que la production de films sonorisés de façon synchrone antérieure aux systèmes développés dans la seconde moitié des années 20 nécessitait, pour des raisons techniques, une fragmentation du spectacle

497. Jean Tedesco, « Pour ou contre le film sonore? », *Cinéa-Ciné pour tous*, N° 114, 1<sup>er</sup> août 1928, cité par E. Arnouldy, *Une histoire du cinéma parlant et sa diagonale mobile. De l'art cinématographique, des « films chantants et parlants » et de leurs séries*, tome 2, thèse de doctorat, Université de Liège, 2001, p. 378.

rétive à l'hégémonie de la mise en forme narrative. On peut en effet suivre l'hypothèse émise par Michel Marie selon laquelle l'essor de la sonorisation des films à l'époque « muette » fut entravé par le succès du long métrage, les courtes bandes susceptibles d'être sonorisées ne pouvant s'intégrer « à un projet narratif dans lequel la parole serait au service de l'histoire racontée »<sup>498</sup>. Les Phonoscènes Gaumont, les *Tonbilder* de Messter ou toutes les bandes réalisées pour des systèmes parents du Chronophone (le Vitaphone, le Synchronoscope, etc.)<sup>499</sup> exhibaient la voix le temps d'une chanson ou d'un sketch, les opéras ou opérettes qui fournissaient parfois le sujet du film étant réduits à un bref extrait. Le recours à des fragments d'œuvres narratives pouvait tout au plus impliquer cette narration « extrinsèque » que Gaudreault identifie dans les films des premiers temps. Francesco Bono remarque que l'augmentation de la longueur moyenne des films constitua en Allemagne un tournant dans la relation du cinéma et de l'opérette : on passe de l'attraction filmée à l'emprunt d'histoires<sup>500</sup>. Les films sonores de Messter, d'une durée d'environ cinq minutes, se limitent au premier stade. L'utilisation éventuelle de plusieurs gramophones en alternance pour augmenter la durée et la cohérence du programme sonore ne changeait pas fondamentalement un principe qui gouvernait déjà les « vues » Lumière ou les « tableaux » de Méliès : celui d'une esthétique fondamentalement discontinue.

Avant que la domination du narratif ne s'instaure, l'attraction sonore s'est nourrie du contexte créé par l'innovation, propice à un rapport particulier avec le public. Que ce terme d'« attraction », pris dans son acception courante, connaisse un tel emploi chez un critique éclairé comme Alexandre Arnoux lorsqu'il discute l'état du cinéma parlant me semble être un indice de sa possible adaptation à la dimension sonore. La mise en évidence de cette terminologie – qui nécessiterait d'être élargie à un plus vaste corpus d'écrits – me semble suggérer un rapprochement entre les deux périodes que sont la première et la quatrième décennie du cinéma, et donc d'un déplacement de la question du « cinéma des attractions » vers l'étude de la voix. Il ne faut bien sûr pas occulter une inévitable

498. Michel Marie, « La carpe et le ventriloque », in Ch. Belaygue (dir.), *Le Passage du muet au parlant*, op. cit., p. 12.

499. Voir Douglas Gomery, « The Coming of the Talkies: Invention, Innovation, and Diffusion », in Tino Balio (dir.), *The American Film Industry*, Madison : University of Wisconsin Press, 1976, p. 195.

500. F. Bono, *art. cit.*, pp. 31-32.

évolution du terme: il ne fait aucun doute que le mot « attraction » ne recoupe pas exactement la même réalité au début du siècle qu'à la fin des années 20 (c'est pourquoi j'ai nettement distingué la voix-attraction du parlant de celle du « cinéma » parlé), et que sa reprise dans le champ de la théorie du cinéma ne se réduit pas, loin s'en faut, à son acception commune, même actuelle.

### V.3 L'ATTRACTION VOCALE DE TYPE « VAUDEVILLE »

Bien que son importance soit relativisée par certains historiens du cinéma<sup>501</sup>, l'exploitation cinématographique intégrée aux théâtres de vaudeville est devenue, notamment grâce aux travaux de Robert C. Allen (dont la thèse de doctorat, publiée en 1980, s'intitule *Vaudeville and Film, 1895-1915. A Study in Media Interaction*), l'un des paramètres importants de l'étude des premières années de l'exploitation cinématographique. Il me paraît intéressant de souligner une certaine filiation entre ces deux séries culturelles qui, comme on le verra, se sont croisées à travers diverses pratiques. Sur un plan socio-économique, on peut noter que le vaudeville participait d'un mode d'élaboration et de consommation du spectacle dont le cinéma – du moins le cinéma hollywoodien – se fera l'héritier tout en le radicalisant, puisqu'il ajoutera au système de distribution à grande échelle mis en place par les entrepreneurs de la fin du XIX<sup>e</sup> siècle une massification opérée dans le domaine de la production<sup>502</sup>. On peut donc dire que les grands promoteurs du vaudeville tels que Keith et Albee ou Marcus Loew initièrent avec leurs circuits une façon de gérer le spectacle qui présente d'évidentes analogies avec la structure à intégration verticale des *majors* de l'industrie cinématographique. C'est pourquoi il n'est pas étonnant de retrouver dans celle-ci ces mêmes promoteurs: le rapprochement entre les groupes Keith-Albee et Orpheum aboutira, après moult transformations, à la constitution de la RKO en 1928, alors que le vaste

501. Voir Ch. Musser, *The Emergence of Cinema, op. cit.*, pp. 297-298. L'auteur met en évidence combien les promoteurs de vaudevilles reléguèrent le film au statut d'un acte de qualité inférieure, ainsi qu'en témoignerait le fait que l'on ne mentionnait pas les projections sur les affiches et que les critiques n'en rendaient pas compte. Cette place marginale accordée au spectacle cinématographique n'interdit toutefois pas que le contexte du vaudeville ait exercé une influence sur la façon dont on montrait et dont on percevait les films.

502. Jeanne T. Allen, « Copyright and Early Theater, Vaudeville, and Film Competition », in John L. Fell (dir.), *Film Before Griffith*, Berkeley; Los Angeles; London: California Press, 1983, p. 78.



circuit que Loew dirige parallèlement à l'exploitation de nickelodéons conduira à la fondation de la MGM. Dans son étude sur la naissance de la culture de masse aux États-Unis, l'historien Jacques Portes est revenu sur la question du vaudeville dans ses liens avec le cinéma et sur les mutations qui ont marqué l'évolution de ce type de spectacle, comme le passage des *concert saloons* aux théâtres de variétés qui permit une diversification de la clientèle, ou l'introduction de capitaux bancaires qui en fit une véritable industrie du spectacle<sup>503</sup>. Au-delà de la circulation des investissements, on peut également comparer vaudeville et cinéma du point de vue du déroulement du spectacle, Allen ayant montré que ce type de divertissement avait donné lieu à une forme stabilisée. La voix constitue à cet égard un point de convergence entre les spectacles scéniques populaires et les projections bonimentées du cinéma des premiers temps, puis le cinéma parlant. Les prestations des bonimenteurs du « muet » et des acteurs des premières années du parlant se sont inévitablement inscrites pour partie dans le modèle spectaculaire de la voix-attraction fourni par le contexte des divertissements de type « vaudeville ». Aux États-Unis, l'une des manifestations les plus courantes de la voix vive à l'époque du « cinéma » parlé résidait dans la pratique de la chanson illustrée qui, justement, représentait une forme courante d'acte de vaudeville avant de se voir intégrée aux nickelodéons<sup>504</sup>. Lorsqu'en 1913 Edison présente sa version améliorée du Kinetophone, c'est avec des magnats du vaudeville qu'il prend contact ; il fait installer son invention dans quatre salles du circuit Keith-Orpheum à New York<sup>505</sup>. Le théâtre de vaudeville semble constituer un lieu où la voix peut s'immiscer dans un ensemble de pratiques spectaculaires diversifiées.

Dans son récent ouvrage sur les « sons du muet », Rick Altman consacre un chapitre aux liens existant entre vaudeville et cinéma, principalement en s'interrogeant sur la place de la musique<sup>506</sup>. À première vue, ses conclusions semblent opposées à ce que j'aimerais démontrer ici, car Altman souligne avant tout combien, durant la première décennie du cinéma, l'exploitation du cinématographe

503. J. Portes, *De la scène à l'écran, op. cit.*

504. Richard Abel, « That Most American of Attractions, the Illustrated Song », in R. Abel, R. Altman (dir.), *The Sounds of Early Cinema, op. cit.*, p. 143.

505. D. Gomery, « The Coming of the Talkies... », *art. cit.*, p. 196.

506. R. Altman, *Silent Film Sound, op. cit.*, chapitre 6.

dans le cadre des spectacles de vaudeville a empêché le développement des projections animées, notamment du point de vue du son. Il note incidemment le fait que les programmes ne faisaient pas mention d'un conférencier, même pour les films de voyage qui, à la même époque, étaient bonimentés dans les nickelodéons, lesquels fournirent à son avis un contexte comparativement très favorable à un déploiement productif des « sons du muet »<sup>507</sup>. Selon lui, le film était donc réduit, dans le « régime répressif » du vaudeville, à faire office d'acte final bon marché<sup>508</sup>. Sans contredire les affirmations de cet historien qui reposent sur une base très documentée, il me semble possible d'envisager quelque peu différemment les liens entre vaudeville et cinéma en mettant l'accent sur la voix-attraction en élargissant la période considérée au-delà de l'ère des nickelodéons (le vaudeville se poursuivant parallèlement au cinéma jusqu'aux premières années du parlant) et en concevant l'expression « vaudeville » dans un sens plus général (il s'agit d'un « mode spectaculaire » recouvrant diverses pratiques voisines). D'ailleurs, l'absence de mention d'un bonimenteur dans les programmes ne constitue pas une preuve de leur absence sur scène, mais reflète uniquement une certaine stratégie promotionnelle. Si les spectacles de vaudeville n'offraient pas de véritable espace d'accueil aux projections cinématographiques qui y étaient reléguées à un rang subalterne, ils n'en ont pas moins été un lieu d'exercice d'une parole adressée à un public quantitativement et sociologiquement large et, littéralement, un lieu où paroles et images filmiques se sont rencontrées (fût-ce de manière différée). Même si personne n'accompagnait les projections de paroles et de musique, ces deux composantes intervenaient nécessairement ailleurs dans la séance, où les films ne constituaient qu'une attraction parmi d'autres. Il me semble que de tels « voisinages » méritent plus d'attention. Ces interactions se sont d'ailleurs poursuivies au-delà de la période du cinéma des premiers temps, l'important promoteur américain M. B. Leavitt affirmant, en 1912, qu'un « grand nombre des théâtres de vaudeville habituels dépendent des bénéfices obtenus grâce aux projections animées »<sup>509</sup>. Notons que cette exploitation du cinéma dans les théâtres de vaudeville continue

507. *Ibid.*, p. 103.

508. *Ibid.*, p. 115.

509. Cité par R.W. Snyder, *The Voice of the City, Vaudeville and Popular Culture in New York*, Oxford; New York: Oxford University Press, 1989, p. 157.

jusque dans les années 20, période où elle devient même la pratique dominante : Snyder fait remarquer qu'en 1928, alors que le son apparaît dans certaines salles de cinéma, seulement quatre établissements de vaudeville sur l'ensemble du territoire américain n'offraient pas de projections cinématographiques<sup>510</sup>.

Compris au sens large, le vaudeville américain peut être rapproché, en dépit de certaines différences que j'évoquerai, des divertissements populaires français comme les cafés-concerts ou le music-hall qui, dès 1903, connurent une forte ingérence du cinématographe puis, quelques années plus tard, accueilleront en dehors de leurs programmes des projections afin de rentabiliser les locaux<sup>511</sup>. Or on peut imaginer que dans un lieu identique, les spectateurs ont tendance à formuler implicitement le même type d'attentes. Cette rencontre entre vaudeville et cinéma s'effectua d'ailleurs aussi dans le sens inverse, puisque Rick Altman précise que les nickelodéons n'étaient pas entièrement consacrés à la diffusion de films, mais proposaient également, entre autres, des actes de vaudeville<sup>512</sup>. En tant que série culturelle qui a coexisté, parfois même spatialement, avec les projections animées, le vaudeville a contribué à instaurer un ensemble d'attentes auprès des spectateurs et d'habitudes chez les exploitants, qui ne furent pas sans incidences sur la diffusion et la réception du médium « cinéma ». Parmi les nombreuses références que Rick Altman fournit dans d'autres chapitres que celui spécifiquement consacré au vaudeville, il apparaît que ces liens sont particulièrement étroits du point de vue de la composante sonore. Altman note par exemple que les premiers films sonores synchronisés, significativement désignés par le terme « actes » (« acts ») dans la presse, visaient à reproduire avant tout la voix chantée afin de procurer à moindres coûts les performances des célébrités du vaudeville ; il remarque également que la plupart des bandes réalisées par la firme Edison pour le Kinetophone « étaient des enregistrements serviles d'actes de vaudeville »<sup>513</sup>. À ce propos, il cite l'un des chroniqueurs du *Motion Picture World* qui écrit en janvier 1913, juste après la première démonstration du Kinetophone, que les exploitants ne cherchent pas à acquérir les droits du nouveau système « parce qu'ils réalisent que la place des images parlantes est et sera

510. *Ibid.*, p. 158.

511. J.-J. Meusy, *op. cit.*, pp. 118 et 137.

512. R. Altman, *Silent Film Sound, op. cit.*, p. 181.

513. *Ibid.*, pp. 162 et 175.

probablement toujours le théâtre de vaudeville... L'image parlante appartient à la scène parlante, tout comme le film muet [silent drama] a plus de succès dans sa propre maison.»<sup>514</sup> On constate que l'opposition entre muet et parlant tend à cloisonner les pratiques, la voix étant associée dans les années 10 aux performances scéniques du vaudeville. Lorsqu'on intégrera la voix aux films, cette association continuera d'opérer au niveau du choix des sujets représentés.

L'utilisation du Kinetophone a donc conduit à une situation d'intermédialité. La projection d'images animées accompagnée de voix synchrones ne s'est pas limitée à l'appareil conçu par Edison ou à des techniques similaires. Il faut en effet prendre en compte une pratique du cinéma parlé qui consistait à recourir à des acteurs qui «doublaient» les personnages de l'écran depuis les coulisses. Cette dernière était maîtrisée par des compagnies qui, précisément, ont connu un essor considérable à l'intérieur de certains circuits de vaudeville comme ceux de Marcus Loew et d'Adolph Zukor<sup>515</sup>. Même les paroles «virtuelles» du film d'animation *Gertie le dinosaure* (1914) de Winsor McCay, un artiste rompu aux pratiques des *dime museum* et des vaudevilles, s'inscrivent pleinement dans le régime spectaculaire de la voix-attraction : le dessinateur demande au monstre de l'écran par le truchement d'intertitres d'effectuer certains mouvements pour le public, puis apparaît lui-même, greffé dans l'image, pour faire une révérence et quitter la scène. En se représentant en montreur de cirque, le fouet à la main – à l'instar d'Ustinov dans *Lola Montès* ou de Chaplin dans le spectacle de puces des *Feux de la rampe* (*Limelight*, 1952) – McCay réfère l'espace diégétique du dessin animé, pourtant totalement virtuel, au modèle de la scène du vaudeville et du music-hall. Par ailleurs, la présentation de Gertie est encadrée, dans la version que diffusa la Vitagraph, par des scènes tournées en prise de vues réelles comprenant des intertitres; or cet encadrement était destiné à remplacer le «boniment» et les performances (les «croquis éclair») dont McCay accompagnait la projection de ses métrages sur la scène des théâtres de vaudeville. Cette intégration dans le «texte filmique» des particularités orales du spectacle de vaudeville témoigne de l'importance des transferts qui s'opéraient entre différentes séries culturelles.

514. *Ibid.*, p. 177.

515. *Ibid.*, p. 169.

Les liens entre le mode spectaculaire du vaudeville et la projection de films accompagnée d'une voix me semblent en outre se resserrer à la fin des années 20. L'influence des spectacles vivants sur les sujets des films parlants est importante mais de brève durée, car contraire aux principes esthétiques de la standardisation qui est en cours. Il faut rappeler que le spectacle de vaudeville qui s'est imposé depuis 1880 disparaît justement avec l'arrivée des *talkies*: le Palace de New York, dernier théâtre en activité, ferme en novembre 1932<sup>516</sup>. On peut donc supposer que les deux formes de spectacle, en concurrence depuis une vingtaine d'années, s'affrontent dès lors sur le même terrain en répondant à des attentes similaires. Lorsque Jacques Portes affirme que la crise économique constitue l'élément essentiel de la disparition du vaudeville en soulignant que «ses salles ont attiré des millions d'Américains alors même que le cinéma constituait un rival redoutable, et [que] seule la grande crise de 1929 a abouti à leur fermeture définitive»<sup>517</sup>, peut-être sous-estime-t-il l'importance de la mutation technique que connaît le cinéma dès 1927-1928 et qui contribue à le préserver quelque peu des effets de la crise. Cet adversaire n'en sera que plus dangereux après l'acquisition de la voix synchrone qui lui permettra de s'approprier le répertoire et les artistes du vaudeville pour mieux le supplanter, ainsi qu'avait déjà commencé à le faire la radio (notamment en engageant des stars de la scène comme Eddie Cantor). Lorsque la Warner commence à produire des courts métrages présentant des sketches ou des pièces en un seul acte (que Charles Wolfe, et à sa suite Édouard Arnouldy, appellent significativement des *vaudfilms*), elle recourt notamment aux services de Brian Foy, qui a fait carrière dans le domaine du vaudeville<sup>518</sup>. Sans tomber dans la vision simpliste d'une stricte causalité, on peut dégager quelques points de convergence entre ces deux séries culturelles du point de vue de l'oralité.

Si l'on part du postulat qu'il existe des croisements entre vaudeville et cinéma parlé ou parlant, on peut rendre compte de la continuité (et non de la «rupture» mythique) et de l'intermédialité introduites par la réalisation et le succès du *Chanteur de jazz*. Le film d'Alan Crosland s'inscrit en effet dans un rapport lâche mais

516. J. Portes, *op. cit.*, p. 98.

517. *Ibid.*, p. 81.

518. Ch. Wolfe, «Vitaphone Shorts...», *art. cit.*, p. 25; É. Arnouldy, *Pour une histoire culturelle...*, pp. 61, 65 et 69.

néanmoins notable à la tradition des *blackface minstrels* (où, comme Al Jolson, des acteurs blancs sont grimés en noir), spectacle très populaire sous sa forme autonome jusque vers 1860 avant de constituer un type de saynète dans les spectacles de vaudeville<sup>519</sup>. En dépit de l'intégration des moments chantés par Al Jolson à l'espace diégétique, ceux-ci sont autant d'attractions ponctuant un film majoritairement muet, à l'image de la succession des numéros dans le vaudeville qui comprend des performances muettes, parlées et chantées. Dans *Le Chanteur de jazz*, le chant s'offre comme une attraction autonome qui ne laisse place à aucune autre manifestation sonore, à l'instar d'un titre musical enregistré sur un disque. De plus, l'unique séquence parlée où Jack s'adresse à sa mère s'apparente à la prestation scénique du monologue, courante dans le vaudeville. Ce moment parlé n'échappe d'ailleurs pas totalement à l'inféodation au musical tel qu'on le pratique dans les spectacles vivants, dans la mesure où il est scandé d'un tempo au piano, comme pourrait l'être une adresse au public entre deux morceaux<sup>520</sup>. Malgré le synchronisme, cet usage du son rappelle l'extradiegéticité (et même l'exogénéité) des accompagnements musicaux de projections auxquels les promoteurs de vaudevilles ne se privaient pas de recourir, puisqu'ils avaient de toute manière les musiciens sous la main<sup>521</sup>.

Revues et vaudevilles offrent aux *talkies* un répertoire non seulement de chansons et de mises en scène, mais aussi d'acteurs comme Eddie Cantor, James Cagney ou Groucho Marx, qui débütèrent dans le circuit Albee<sup>522</sup>. Les firmes hollywoodiennes viendront y puiser des techniques et des savoir-faire pour réaliser bon nombre des premières productions parlantes. Il ne s'agit bien sûr pas de réactiver le *topos* du « théâtre filmé » propre aux réfractaires de l'époque, mais d'émettre l'hypothèse selon laquelle la vigueur de l'implantation du vaudeville et de sa tradition culturelle a favorisé aux États-Unis l'intérêt massif et rapide du public et des dirigeants de l'industrie cinématographique pour le parlant. Car, même si l'histoire économique et des techniques permet d'expliquer en partie en quoi cette période a connu la conjonction d'éléments favorables au

519. Voir J. Portes, *op. cit.*, pp. 29-31 et 46.

520. Voir Youssef Ishaghpour, « La parole et l'interdit des images : *Le Chanteur de jazz* et les juifs d'Hollywood » [1999], in *Historicité du cinéma*, Tours : Farrago, 2004, p. 94.

521. J. Portes, *op. cit.*, p. 188.

522. *Ibid.*, p. 93.

parlant, les raisons de l'absence de véritable percée commerciale des Phonoscènes ou des « Film parlants » réalisés à partir de 1906 sous l'impulsion du pionnier français Léon Gaumont restent passablement obscures, les imperfections techniques de ce système, soulignées par Gomery<sup>523</sup>, n'étant probablement pas seules en cause. Cette absence d'institutionnalisation vaut d'ailleurs pour tous les autres pays producteurs de films. On a évoqué l'Allemagne avec Messter, mais la situation est identique au Québec, Gaudreault et Sirois-Trahan notant qu'en 1907 « Montréal fut [...] le terrain d'une véritable course contre la montre pour doter les salles d'une « machine parlante », une machine qui ne réussit pas, à l'époque, à sortir du cadre étroit des représentations spéciales »<sup>524</sup>.

La filiation directe que l'on trouve aux États-Unis entre le cinéma parlant et les prémices d'une culture de masse pourrait avoir contribué à mettre plus ostensiblement en lumière le potentiel commercial de l'invention, et prédisposé le public de façon plus profonde à assister à un spectacle dont la structure et les conditions de réception relèvent d'une dominante « attractionnelle », tandis que les tentatives de Gaumont, stoppées par la Première Guerre mondiale, sont quant à elles contemporaines d'une période cruciale en termes d'affirmation de la structure narrative durant laquelle les « attractions sonores » ne sont plus en adéquation avec les nouvelles exigences du public. En effet, pour saisir l'évolution qu'a connue le cinéma durant la période d'exploitation du Chronophone – on peut prendre comme points de repère l'installation de ce dispositif au Théâtre du Gymnase en été 1908, puis au Gaumont Palace dès 1911, il suffit de penser que, dans la production significative de D. W. Griffith, l'utilisation de cet appareil de sonorisation se situe entre *The Adventures of Dollie* (1908) et *The Lonedale Operator* (1911), c'est-à-dire dans une phase de rapide évolution des pratiques filmiques qui voit se généraliser le montage alterné, le gros plan, les raccords basés sur des « sons virtuels », etc. dans le cadre de films reposant sur le canevas de la course-poursuite. Le développement de ces pratiques, qui se sont imposées au début de la Première Guerre mondiale, ne sera pas sans incidences sur la longueur des métrages et sur la complexité des récits contés. Dans son étude

523. D. Gomery, *op. cit.*, p. 195.

524. André Gaudreault et Jean-Pierre Sirois-Trahan, « La bataille du parlant, rue Sainte-Catherine, vingt ans avant le *Jazz Singer* », in A. Gaudreault (dir.), *Au pays des ennemis du cinéma...*, *op. cit.*, p. 147.

de l'histoire du « scénario » au cinéma, Alexander Schwarz constate que, pour la période 1907-1913, les versions écrites de films publiées en Allemagne et en Russie étaient essentiellement des textes attractionnels (« *Attractionstexte* »), et donc continuaient d'obéir à l'hétérogénéité du spectacle initiée à l'ère du vaudeville (« *Vaudevillezeit* »)<sup>525</sup>. La consolidation du MRI au sens de Burch s'est par contre sans conteste accompagnée d'une domination de l'*écrit*.

La prompt internationale du succès du parlant témoigne bien sûr d'un intérêt similaire dans d'autres pays – le music-hall en France ou l'opérette en Allemagne jouent un rôle comparable au vaudeville américain – mais n'invalide pas le fait que le vaudeville a constitué aux États-Unis un ferment propice à la réception des *talkies*. En examinant le vaudeville sur la base des écrits de certains commentateurs, j'entends dégager des éléments qui ont trait à la plupart des arts de la scène qui lui sont contemporains, pour autant qu'ils se situent du côté du spectacle populaire, et non du théâtre légitimé.

Le sens le plus ancien du terme d'origine française « vaudeville » renvoie à un type de chanson populaire. Cette notion semble donc avoir l'oralité si bien chevillée au corps que l'on serait tenté de la rapporter étymologiquement, comme le fait Robert W. Snyder en intitulant métaphoriquement son ouvrage sur la question *Voice of the City* (1989), non pas à la référence attestée à un espace géographique (le Vau de Vire), mais à la présence de la voix. Bien que le mot « voix » ne puisse être un produit phonétique de « vau », le glissement semble s'être imposé dans un usage ultérieur. Le terme qualifie alors un spectacle qui comprend des chants et qui, il est vrai, s'implanta principalement dans les grands centres urbains. Ce type de spectacle s'est développé aux USA, alors qu'en France la même notion prenait un autre sens pour se référer au ton des histoires racontées, celui de la comédie légère, et non au mode de présentation (on parlera plutôt de « café-concert » pour évoquer une pratique parente du vaudeville américain). Dans l'acception française du terme, le vaudeville a évolué vers une diminution de l'hétérogénéité du spectacle et vers une domination du langage parlé: les « situations comiques entrecoupées d'airs connus » se voient bientôt privées de toute musique<sup>526</sup>. En Amérique par

525. A. Schwarz, *Der Geschriebene Film...*, *op. cit.*, pp. 50-51.

526. J. Portes, *op. cit.*, p. 79.



contre, le chant constitue un des éléments définitoires de cette pratique. L'accent y est donc mis sur la voix elle-même tout autant que sur le référent du texte.

Lorsque Jeanne T. Allen évoque les problèmes que rencontrèrent les promoteurs de vaudeville pour protéger leurs actes par un copyright, elle met en évidence les particularités de ce type de spectacle, notamment l'absence de notations écrites et le faible degré de narrativité des numéros. Les indécisions dont témoignent selon elle les instances juridiques font écho, en raison de la difficulté qu'elles rencontrent à saisir ce qui est spécifique au moyen d'expression, à celles qu'André Gaudreault a relevées dans les documents relatifs aux poursuites judiciaires entreprises par les différentes compagnies cinématographiques américaines au tournant du siècle<sup>527</sup>. Allen note par exemple qu'« en 1914 encore, les cours déclaraient que la voix, les mouvements et les postures d'un acteur ainsi que la pure gestion scénique ne possédaient aucune qualité littéraire et ne pouvaient être protégés par le copyright théâtral »<sup>528</sup>. Parmi les éléments mobilisés pour la création d'une performance scénique de vaudeville, la voix occupe la première place en tant que manifestation pleine de l'*oralité*. Elle s'y distingue du texte, s'autonomise comme facteur primordial de l'expression. Les diverses modalités vocales sont néanmoins liées au déroulement global du spectacle dont certains principes généraux me paraissent consonner avec les thèses zumthoriennes que j'ai appliquées à la période du « cinéma » parlé.

Récemment, Jean Mottet a proposé une analyse pertinente de la forme spectaculaire du vaudeville. Il déclare de prime abord que, « dans leur succession, les *acts* du *vaudeville* échappent à la mécanique du récit », même s'ils obéissent individuellement à la structure d'un « micro-récit »<sup>529</sup>. Ce rapport de la partie au tout s'apparente au « cinéma des attractions » dont j'ai dit que les bandes n'étaient pas juxtaposées selon une logique narrative, bien qu'elles présentassent selon les cas un certain degré de narrativité. L'agencement des actes correspond en fait à l'organisation filmographique des « plans » des premiers films. Cette mise en retrait de la

527. A. Gaudreault, *Du littéraire...*, *op. cit.*, pp. 121-131.

528. J. T. Allen, « Copyright and Early Theater... », *art. cit.*, p. 183.

529. Jean Mottet, « Le spectacle de *vaudeville*: esthétique de la disjonction, ancrage dans le présent et stéréotype », in Claudine Dupré La Tour, André Gaudreault, Roberta Pearson (dir.), *Le Cinéma au tournant du siècle*, Lausanne; Québec: Payot; Nota bene, 1999, p. 152.

narration distingue fortement le vaudeville des œuvres littéraires, et influe sur la représentation du temps qui, n'étant pas vectorisé comme dans le roman par la finalité de certaines actions, est perçu comme relevant exclusivement du présent. Mottet avance deux hypothèses pour expliquer ce sentiment de présentification qui constitue selon moi l'un des traits caractéristiques de la voix-attraction : d'une part, la mise en scène est dense au niveau de la représentation visuelle (ou audiovisuelle, si l'on inclut bruit, musique et voix indépendamment du verbe), alors que le texte y est pauvre ; d'autre part, il y a dans ce type de spectacle « absence de voix narrative », alors que « derrière le dialogue dramatique [du théâtre légitimé], un effet d'auteur persistant permet au spectateur de rattacher la réception de l'ensemble au discours de quelqu'un »<sup>530</sup>. Le premier aspect a trait à l'espace de la performance tel que le conçoit Paul Zumthor. Le second se situe au niveau de l'énonciation narrative : on peut comprendre l'« effet d'auteur » comme la « présence de l'énonciateur », même si le terme « auteur » prend tout son sens dans le contexte juridique concret évoqué par Jeanne T. Allen. Alors que Gaudreault a montré que la décision d'un juge fit jurisprudence en 1905 en permettant aux vues constituant un « film » d'être enregistrées comme un ensemble au copyright<sup>531</sup>, cette reconnaissance d'un « narrateur » qui unifie le discours ne semble pas avoir été possible dans le domaine du vaudeville, ainsi qu'en témoignent les documents analysés par Jeanne T. Allen. C'est ce dont Mottet rend compte sur le plan théorique en opposant l'oral (la performance d'un acteur dans le présent de la représentation) à l'écrit (le texte d'un auteur), l'éclatement du spectacle à l'unité de l'œuvre. Or, il est intéressant que Jean Mottet utilise la notion de « voix narrative » qu'il associe à la marque d'un auteur et au rejet dans le passé des actions montrées (je reviendrai sur cette question en VI.3). Dans le vaudeville, l'instance énonciatrice s'incarne dans le *showman* présent sur scène, non dans une instance de création absente au moment où l'œuvre est consommée ; la « voix narrative » abstraite est supplantée par les voix audibles d'acteurs visibles. La distinction opérée par Mottet sur le plan de la temporalité entre le vaudeville et le théâtre rejoint les propos de Sacha Guitry lorsqu'il souligne la différence entre le théâtre et le

530. *Ibid.*, pp. 155-156.

531. A. Gaudreault, *Du littéraire...*, *op. cit.*, p. 130.

cinéma à l'époque de la réalisation du *Roman d'un tricheur*: selon Guitry, l'acteur de cinéma «a joué», alors que le comédien sur scène «joue» (voir chapitre III). Pour Mottet, qui creuse encore plus l'écart entre les arts de la scène et le cinéma parlant, les événements du théâtre sont eux-mêmes perçus comme passés comparativement aux actes du vaudeville. Ses remarques peuvent partiellement s'appliquer au «cinéma» parlé, dont la représentation visuelle est également densifiée par le mélange de performances *live* et d'images projetées. Toutefois, le bonimenteur introduit dans ce spectacle une «voix narrative» à travers tout ce qui relève, dans son activité, de l'établissement d'une cohérence entre les différentes vues. Mottet exagère par contre l'absence d'individualisation de la voix du *showman*: la voix-attraction est elle aussi ancrée dans une instance spécifique, quand bien même l'ensemble du spectacle ne se donne pas comme le produit de cette voix. La voix-attraction étant fortement liée à la prestation individuelle du *showman*, elle s'oppose au contrôle d'une instance invisible, désincarnée, comme l'est la voix-over dans le cinéma parlant.

On peut se demander à quoi ressemblait la voix-attraction à laquelle recouraient les acteurs du vaudeville. Mottet a montré que ce type de spectacle utilise des représentations stéréotypées pour se rapprocher du public avec lequel il «recherche la proximité, l'intimité»<sup>532</sup>, notamment en optant pour une forme interpellative. On retrouve donc le mode de l'adresse qui, pour Gaudreault et Gunning, définit le régime spectaculaire particulier du «cinéma des attractions». Eisenstein se référait déjà aux spectacles de type «vaudeville» lorsqu'il parlait des attractions cinématographiques en qualifiant «le music-hall et le cirque» d'«école du monteur». Le théoricien soviétique soulignait également combien, du point de vue attractionnel, paroles et voix sont mises sur un pied d'égalité avec les éléments de la représentation visuelle: «Pour le spectateur, la «voix d'or» d'Ostoujev ne vaut pas plus que la couleur du maillot de la prima donna, le coup de timbale vaut autant que le monologue de Roméo.»<sup>533</sup> Il s'agit de renverser la hiérarchie que l'on trouve par exemple dans l'opéra et le théâtre légitimé entre la voix des acteurs qui est porteuse du texte ou de la mélodie composée par un «auteur», et toutes les composantes non verbales du

532. J. Mottet, «Le spectacle de vaudeville...», *art. cit.*, p. 162.

533. S. M. Eisenstein, «Le montage des attractions», *art. cit.*, pp. 117 et 119.

spectacle susceptibles d'établir un rapport d'immédiateté avec le public. Non pas que la voix soit évincée de l'esthétique des attractions : elle s'y mêle à quantité d'autres stimuli audiovisuels.

Comme le montre la description de la structure idéale d'un spectacle de vaudeville que donne George Gottlieb dans un ouvrage intitulé *Writing for Vaudeville* (1916), le principe central de ce type de représentation réside dans la stimulation du public qu'il s'agit de réveiller, de maintenir dans l'expectative et de soumettre à une tension croissante<sup>534</sup>. Si l'ordonnance du *show*, basée sur une juxtaposition des parties, y est essentielle, elle est pensée en termes d'effet sur le spectateur. Étant le moyen le plus courant de communication, la voix représente l'un des paramètres principaux de l'établissement d'un lien avec le public. Le vaudeville se caractérise par cette même capacité d'adaptation au public visé que le cinéma bonimenté, dont Germain Lacasse a montré qu'il participait d'une « pratique de résistance ». Snyder souligne la nécessité pour les promoteurs de vaudeville de prendre en compte les divers contextes socioculturels dont étaient issus leurs spectateurs, précisant qu'« il était impossible d'intégrer totalement la population de la ville dans quelques théâtres standardisés »<sup>535</sup>. Cette adaptation conduit à privilégier une communication immédiate et empathique entre acteurs et spectateurs. Snyder illustre cette adresse directe et littéralement « attractionnelle » (« *direct appeal* ») en citant un article du *New Republic* paru en 1922 dans lequel l'auteur, Mary Cass Canfield, décrit en quoi consiste généralement la performance d'un acteur de vaudeville :

« Il plaisante avec le chef d'orchestre, il confie à ses auditeurs des histoires inventées à propos de la direction de l'établissement, d'autres acteurs, à propos du monde captivant des coulisses ; il s'adresse à des complices placés au troisième rang ou au balcon et fait en sorte de se moquer d'eux pour la joie de tous. Il enjoint les [...] auditeurs à participer au refrain de sa chanson ; ils obéissent par des clameurs. Le public devient une partie du *show* et aime l'être. Il y a de l'art communautaire pour toi. »<sup>536</sup>

534. Cité par R. W. Snyder, *op. cit.*, pp. 66-67.

535. *Ibid.*, p. 91.

536. *Ibid.*, pp. 104-105.

Par la parole, le comédien met en relation de façon réflexive toutes les instances de «l'espace de la production» et assure une médiation entre le public et le spectacle. La «scène» s'élargit dès lors à l'espace complet de la salle, incitant le public à se considérer comme un «acteur» de la représentation. Le chant, comme l'humour, joue un rôle fédérateur. Jacques Portes relève combien, dans les spectacles populaires de la fin du XIX<sup>e</sup> siècle, les chansons appartiennent à un fonds culturel commun et sont exploitées pour créer une participation active du public :

«Les spectateurs se pressent, bissent, huent ou s'exclament. Un certain nombre d'airs sont tellement connus du public qu'ils sont intégralement repris avec enthousiasme, quand ils ne sont pas demandés au chanteur ou au musicien, même s'ils ne font pas partie du programme.»<sup>537</sup>

Cette attitude n'est pas sans rappeler la manifestation plus explicitement politique du public de la mémorable séance de cinéma bonimenté qu'évoque Rabon (voir II.1), durant laquelle les spectateurs exigèrent que les musiciens jouent l'hymne français.

Le régime d'oralité est induit par la structure du vaudeville, dont le caractère *discontinu* interdit l'effacement durable de l'espace-temps de la salle de spectacle devant celui de l'univers diégétique. À la «fixation» de la représentation de certains spectacles (par l'écriture, l'inscription phonographique et photographique), le vaudeville oppose une modularité extrême: «Rien n'est fixé définitivement, la seule règle étant la régularité des séquences.»<sup>538</sup> Les variations s'opèrent autant d'un acte à l'autre – selon J. T. Allen, l'acte constitue «une unité discrète interchangeable dans un système de huit ou neuf actes»<sup>539</sup> – qu'au sein même des actes grâce à l'improvisation du *showman* et aux réactions spontanées du public. Le vaudeville est par excellence le lieu d'exercice de la voix vive selon le régime de la voix-attraction.

D'autres spectacles populaires, tels que le music-hall ou le café-concert, partagent avec le vaudeville américain les mêmes caractéristiques. Certes, on y trouve moins de diversité et d'hétérogénéité, le music-hall étant plus fortement centré sur une vedette unique et

537. J. Portes, *op. cit.*, p. 34.

538. *Ibid.*, p. 84.

539. J. T. Allen, «Copyright and Early Theater...», *art. cit.*, p. 181.

le café-concert sur la seule chanson<sup>540</sup>. Si le vaudeville a constitué l'un des modèles de référence du cinéma parlant développé aux États-Unis, ces spectacles scéniques ont également connu en France des liens avec les technologies du son puisqu'ils ont fait l'objet des premières transpositions phonographiques. Dans son examen du catalogue de l'éditeur Pathé de 1898, Giusy Pisano a constaté que les sujets des enregistrements qui n'étaient pas des emprunts au répertoire traditionnel provenaient « du café-concert d'abord et du music-hall ensuite », comme en témoigne la classification utilisée à l'époque (« pochardes », « comique-troupiers », « patriotique », etc.)<sup>541</sup>. Les chansons et monologues éprouvés sur scène s'immiscèrent donc nécessairement dans les expérimentations de couplage du phonographe et du cinématographe. Les cylindres pré-enregistrés offraient un sujet aux bandes filmiques qui, généralement tournées en *play-back*, se prêtaient bien à l'illustration de chants préexistants.

Dans les années 50, Roland Barthes s'est penché sur la question de la spécificité de ce type de spectacle dans un texte intitulé « Au music-hall » en mettant l'accent sur l'importance du « geste », qu'il discute en des termes qu'il nous semble possible d'appliquer à la voix (on constate d'ailleurs dans l'exemple de Rabon que l'incitation verbale mène aisément à l'action physique). L'écrivain a conscience que le music-hall français se distingue de son « équivalent » américain par son absence de chant. On pourrait ajouter qu'il se caractérise également par une mise en retrait de la parole, celle-ci étant traditionnellement réservée en France à un mode de représentation plus légitimé (le théâtre). Le film *Les Enfants du Paradis* (Marcel Carné, 1944) dépeint cette époque où l'interdiction de recourir au langage parlé pesait sur de tels établissements. Elle disparaît certes en 1867 avec la levée des restrictions destinées à protéger les théâtres<sup>542</sup>, mais une solide tradition de mime, d'acrobatie et de chant s'est établie, qui se poursuivra. L'opposition entre des arts nobles parlants et un divertissement populaire muet ou chantant constitue donc, au tournant du siècle, un paradigme important au sein de l'espace culturel français qui n'a pas connu le « mélodrame » (au sens d'un spectacle associant le langage parlé et

540. Selon J. Portes, *op. cit.*, pp. 82 et 85.

541. Giusy Pisano, *Interférences entre son et image : un siècle de recherches avant le cinéma parlant*, thèse de doctorat, Paris, 1999, p. 277.

542. J.-J. Meusy, *op. cit.*, p. 114.

des passages musicaux) qui marqua le public américain au cours du XIX<sup>e</sup> siècle<sup>543</sup>. On retrouve dans les considérations de Barthes sur le music-hall une attention à cette fonction de présentification qui privilégie l'espace-temps scénique sur la constitution d'une temporalité diégétique. Entre le geste et la temporalité du déroulement spectaculaire s'établit un rapport étroit d'immédiateté qui instaure un temps que Barthes qualifie de «juste, sidéral, le temps de la chose elle-même». Le geste de l'acteur de music-hall se distingue selon lui par le fait qu'il est possible de le «dégager de toute cause, [...] [de l'] épuiser comme spectacle et non comme signification»<sup>544</sup>, c'est-à-dire qu'il relève de cette pure spectacularité propre au pôle attractionnel.

Lorsque les maisons de production devront engager, à l'époque de la transition au parlant, des acteurs issus de tels spectacles vivants, elles devront prendre la mesure des différences qui résident entre ces voix-attraction et les voix-action du théâtre traditionnel. James Lastra discute la prise de conscience dont témoignent de nombreux techniciens du son à Hollywood envers l'impasse que signifierait l'adoption d'un unique modèle de «réalisme» sonore. Il cite à ce propos un ingénieur de la MGM, Wesley C. Miller, qui constate en 1929, dans un article de la revue *American Cinematographer*, qu'un acteur de vaudeville peut être enregistré plus «intimement» et donc avec un niveau d'intensité supérieur à celui d'autres styles d'acteur «à cause de la nature de son numéro»<sup>545</sup>. La proximité du spectacle scénique est alors reconduite *via* les caractéristiques phonographiques du cinéma parlant qui, dans les multiples passages chantés de ses premières productions, s'inspire passablement du vaudeville. Miller reconnaît le statut spécifique du mode de l'adresse propre à cette voix-attraction, indice d'une intermédialité qui ne survivra pas à la standardisation des pratiques sonores.

#### **V.4 L'ESTHETIQUE DU VAUDEVILLE DANS LES PROLOGUES DU «CINEMA» PARLE**

L'intégration de la «voix du vaudeville» dans le spectacle cinématographique est manifeste dès les prologues que certains exploitants organisèrent à partir du milieu des années 10 jusqu'à l'interrègne

543. R. Altman, *Silent Film Sound*, *op. cit.*, p. 36.

544. Roland Barthes, «Au music-hall» [1957], repris in *Mythologies*, Paris: Seuil, 1970, pp. 164-165.

545. Cité par J. Lastra, *op. cit.*, p. 173.

du parlant. Ces spectacles fastueux créaient une atmosphère dont la fonction consistait à annoncer les caractéristiques de l'univers diégétique des projections muettes qui leur faisaient suite<sup>546</sup>. Si cette pratique mixte alliant projection et performance *live* ne connut une véritable extension qu'aux États-Unis, c'est probablement, à nouveau, en raison du fort ancrage qu'a connu le modèle du vaudeville dans la culture américaine: Vinzenz Hediger rapproche à plusieurs reprises prologue et vaudeville, et Rudmer Canjels décrit le prologue comme «une continuation de la relation de proximité entre la scène et l'écran, une relation qui avait déjà commencé avec le vaudeville»<sup>547</sup>. Le prologue est un lieu de manifestation de la voix vive, et appartient plus largement au régime de l'oralité dans la mesure où l'«auteur» du spectacle est avant tout l'exploitant qui, à l'instar du promoteur de vaudeville, compose son programme. Alors que certaines maisons de production du début du siècle éditaient des textes destinés aux bonimenteurs pour contrôler quelque peu leurs commentaires, les *press books* de la fin des années 10 contenaient parfois des suggestions de prologues pour éviter que ceux-ci ne dépareillaient trop avec le film<sup>548</sup>.

On sait l'importance qu'occupe le verbal dans les génériques de films. Il me paraît donc intéressant d'aborder une pratique qui fait reculer ce «seuil» en recourant notamment à la voix-attraction, alors que les débuts de films constitueront plus tard le lieu privilégié d'apparition de la voix-narration *over*, celle-ci donnant d'entrée de jeu l'impression de faire naître le récit en occupant le «poste de commande» de l'énonciation filmique. Entre le mode *présentationnel* du prologue qui instaure nécessairement une certaine discontinuité avec le récit filmique et l'immersion diégétique immédiate qu'induit la fonction *représentationnelle* de la plupart des voix-over se distribuent les différents degrés de la voix-narration. La question du prologue permet en outre de dégager un point commun aux deux régimes vocaux antagonistes de la voix-attraction et de la

546. Vinzenz Hediger, «Putting the Spectators in a Receptive Mood», in Veronica Innocenti, Valentina Re (dir.), *Liminal/Le Soglie del film*, Udine: Università degli studi di Udine, 2004. L'article de Hediger, qui se base sur un dépouillement d'une partie de la presse corporative américaine, fournira la plupart des informations mentionnées ici. Je remercie son auteur pour les documents qu'il m'a gracieusement fournis.

547. Rudmer Canjels, «Featuring on Stage: American Prologues from the 1920s», in *Liminal/Le Soglie del film*, *op. cit.*, p. 309.

548. *Ibid.*, p. 313.



voix-narration: la *fonction instauratrice du logos*. «Au commencement était le verbe», nous dit le plus célèbre de tous les récits: il en est du monde comme des univers diégétiques, fussent-ils prioritairement visuels.

La création de prologues présente un intérêt particulier dans le cadre d'un rapprochement entre le cinéma des premiers temps et le cinéma ultérieur, car cette façon de concevoir la séance révèle d'une part, selon Hediger, «une ligne de continuité évidente entre le cinéma primitif et la première période des longs métrages»<sup>549</sup> et, d'autre part, certaines similitudes entre les pratiques du cinéma parlé et l'introduction de la voix synchrone aux débuts du parlant. Sur ce second point, Valentin Mandelstamm, commentateur perspicace de la généralisation du parlant dont il est le témoin, propose en 1929, dans l'un de ses articles de la revue *Cinémagazine*, un historique dans lequel il fait allusion à la filiation existant entre les prologues parlés et le cinéma parlant. Mandelstamm explique ainsi l'adoption du cinéma parlant par l'industrie américaine:

«Les recettes des théâtres de cinéma aux États-Unis, qui depuis 1914 avaient suivi une marche ascendante, se stabilisant vers 1924, s'étaient mises à baisser dans d'inquiétantes proportions; [...] il fallait maintenant les [les superproductions] «encadrer», et, pour cela, dépenser des sommes énormes en publicité, et en présentations coûteuses, comprenant des prologues, des revues, des ballets, des numéros choisis de music-hall; et le film, de la sorte, finissait par passer au *second plan*. [...] Et l'on se demandait où l'on allait s'échouer, lorsque le film synchronisé et le film parlant sont venus sauver la situation.»<sup>550</sup>

Si l'analyse économique de Mandelstamm paraît erronée (puisque l'ajout de performances *live* luxueuses participe bien avant le milieu des années 20 de la stratégie commerciale des exploitants américains), la façon dont il présente la nécessité d'opérer une transition au parlant est significative: pour lui, le parlant permet de remplacer certaines pratiques issues des spectacles scéniques qui tendaient à minorer la place de l'objet «film» dans le programme. Les sons pseudo-attractionnels du film se sont substitués à l'univers

549. V. Hediger, «Putting the Spectators...», *art. cit.*, p. 293.

550. Valentin Mandelstamm, «L'Avènement des films parlants et synchronisés», *Cinémagazine*, N° 5, 1<sup>er</sup> février 1929, p. 194.

sonore des attractions. En dotant le cinéma du son synchrone, «l'encadrement» a pu être intégré au film même, dans un mouvement de fixation des éléments du spectacle auquel différentes évolutions techniques ont contribué.

Le cas le plus manifeste de l'appropriation par le film de ce qui constituait bien souvent une parole-attraction «externe» est celui des prologues filmés, car ceux-ci conservent tout de même, en dépit de leur fonction d'annonce, une certaine autonomie par rapport au long métrage et au récit filmique. La mise à l'écart du prologue est particulièrement grande lorsqu'il consiste en un court métrage réalisé indépendamment du film. Or les premiers *Vitaphone shorts* qui témoignèrent de l'aboutissement de la technologie sur laquelle avait misé la Warner étaient principalement conçus comme des substituts peu coûteux aux prologues *live*<sup>551</sup>. Pensées comme des attractions sonores censées précéder le long métrage muet, ces premières bandes transposaient sur le support filmique les numéros de vaudeville. Les courts métrages ont donc fait office de transition entre prologues *live* et prologues filmiques, puis se sont transformés en introductions totalement solidaires du film. Les marges du film connaissent donc une forme de persistance de la voix-attraction. Avant d'en venir à cette intégration (de premier niveau) de la parole dans le film, on peut s'interroger sur la place accordée au verbal dans les prologues du «muet».

Comme l'a montré Vinzenz Hediger en se concentrant sur la période 1915-1930 qui débute avec la disparition des nickel-odéons et l'avènement de réseaux de salles prestigieuses, le film, attraction principale du spectacle «cinématographique», était valorisé par des performances qui visaient à préparer le spectateur, à le «chauffer» en créant une atmosphère adéquate. Selon la périodisation proposée par Hediger, le début des années 20 est caractérisé par une certaine autonomie des prologues constitués d'actes de vaudeville et de variétés. Cette solution était économiquement plus viable puisqu'elle n'exigeait pas la création d'un spectacle spécifique pour chaque film, les productions scéniques étant destinées à des tournées à l'intérieur d'un circuit de salles<sup>552</sup>. Cette volonté

551. Voir R. Canjels, *art. cit.*, p. 314; Ch. Wolfe, «Vitaphone Shorts...», *art. cit.*, pp. 61 et 67, et «On the Track of the Vitaphone Short», in M. Bandy (dir.), *The Dawn of Sound*, New York: MoMA, 1989, repris et traduit en italien («Sulle tracce dei cortometraggi Vitaphone») in R. Altman (dir.), *L'Immagine acustica II. Il Passaggio dal muto al sonoro in America* (Cinegrafie, N° 6), Ancona: Transeuropea, 1993, p. 22.

552. V. Hediger, «Putting the Spectators...», *art. cit.*, pp. 292 et 295.

de ne pas faire interagir les diverses performances *live* et le film ont eu pour corollaire une tendance à l'exclusion du langage verbal dans les pratiques scéniques. Peut-être cette raréfaction de la parole s'explique-t-elle par la crainte que les prologues ne concurrencent le long métrage en soulignant par contraste ce « manque » qu'est la mutité du film. Rudmer Canjels explique en effet que le soin apporté aux transitions ménagées entre la scène et l'écran résultait de la nécessité d'obtenir une analogie entre les deux représentations, quand bien même le second médium « est dépourvu de son (et surtout de la voix humaine), de couleurs et de tridimensionnalité »<sup>553</sup>. Siegfried Kracauer, qui accorde dans ses critiques de la *Frankfurter Zeitung* une attention particulière aux spectacles vivants accompagnant les films (et plus globalement à l'insertion du film dans la culture de masse de son époque), considère dans son célèbre texte « Le culte de la distraction » que ces « représentations globales, où le film s'insère comme une partie dans un plus vaste ensemble », constituent « l'attraction décisive » des divertissements berlinois<sup>554</sup>. Il reproche toutefois aux exploitants des palaces leur volonté de créer une unité artistique au lieu de jouer de l'éclatement et d'intensifier des effets qu'il juge en totale adéquation avec la vie urbaine moderne. Il estime également que si le cinéma perd de son impact, c'est en raison du voisinage de spectacles vivants qui nuisent grandement à sa « force d'illusion »<sup>555</sup>.

Au sein des diverses pratiques de type « vaudeville » présentées dans certaines séances de « cinéma » dans les années 20 (cirque, défilés de mode, moments chantés ou dansés, pantomimes, numéros d'acrobatie, etc.), le verbal parlé ne jouait pas un rôle primordial. On trouve par contre des exceptions sous l'appellation plus spécifique de « présentation » ou, justement, de « prologue », un terme qui suggère des liens plus étroits avec l'univers des films. Lors de son dépouillement de la publication corporative *Exhibitor's Herald* (devenu *Exhibitor's Herald and Moving Picture World* à l'époque du parlant), Vinzenz Hediger a notamment relevé un article représentatif datant du 16 avril 1921 et signé du département « exploitation » de la First National Pictures. On y propose

553. R. Canjels, *art. cit.*, p. 311.

554. Siegfried Kracauer, « Kult der Zerstreung. Über die Berliner Lichtspielhäuser », *Frankfurter Zeitung*, 4 mars 1926, repris et traduit (« Culte de la distraction ») dans *Le Voyage et la danse: figures de ville et vues de films*, Saint-Denis: Presses Universitaires de Vincennes, 1996, p. 58.

555. *Ibid.*, p. 62.

quatre possibilités de prologues pour la sortie du film de Marshall Neilan *Bob Hampton of Placer*<sup>556</sup>. Ce texte témoigne de l'intérêt porté par une maison de production aux présentations scéniques d'un film dont le potentiel commercial et publicitaire est important. Ces « suggestions » doivent aussi être comprises comme une volonté d'ingérence de la production dans le secteur de l'exploitation, c'est-à-dire comme une tentative de standardisation de cette « attraction » (terme utilisé dans l'article). Dans l'introduction de son texte, l'auteur met en évidence le fait que les quatre propositions peuvent être adaptées à tous types d'établissements, qu'il s'agisse de palaces métropolitains ou de salles de petites villes, enjoignant ainsi à étendre cette pratique au-delà de représentations exceptionnelles dans des salles possédant une scène particulièrement spacieuse. Les versions sont énumérées dans l'ordre croissant de leur complexité en termes d'organisation scénique. Les actions représentées et le décor sont toujours issus du monde de *Bob Hampton of Placer* dont ils offrent une version simplifiée, notamment de l'ouverture du film (version 2) et de sa clôture (version 4), comme si le spectacle sur scène devait répéter les bornes filmiques pour faciliter la continuité. Il est important de noter que, à l'exception de la quatrième version qualifiée de « prologue spectaculaire », aucune d'entre elles ne comprend de langage parlé. La pantomime y est accompagnée de chants (chanson sentimentale entre un homme et une femme dans la première, danse guerrière d'Indiens dans la troisième) et d'effets sonores renvoyant à une action hors-champ (cris, tirs de fusil, hurrahs de soldats et bruits de bottes suggérant leur départ) ou convoyant même une dimension narrative (la progression du sauvetage des immigrants par les soldats est indiquée par l'augmentation du bruit de mitraille). Dans la seconde proposition dénommée « prologue d'action », le caractère superflu du texte est même souligné (« *no dialog is necessary* »). La dernière version est quant à elle plus proche d'une représentation théâtrale puisqu'elle présente un dialogue entre deux personnages : les acteurs ne s'adressent pas au public, mais créent une saynète autonome. En participant ainsi à la mise en place de la diégèse du film, les paroles relèvent majoritairement de la voix-action. Comme il s'agit d'exposer un contexte historique précis (il

556. « First National Exploitation Department Suggests Prologues for New Neilan Film », *Exhibitor's Herald*, Vol. XII, N° 16, 16 avril, 1921, pp. 54-55.

est question du Général Custer et de Sitting Bull), la situation dialoguée naturalise la transmission d'informations vers le public du film. En dépit de la présence de la composante verbale, il est toutefois significatif que l'auteur de l'article ne propose aucun texte pour les personnages, se contentant de résumer brièvement son contenu, alors qu'il décrit avec attention l'action et les postures inspirées d'une peinture célèbre («Custer's Last Stand»). La référence à une œuvre picturale témoigne de l'importance accordée aux effets visuels de la mise en scène et aux liens intertextuels qui créent une complicité avec le spectateur et confèrent à la scène une valeur ajoutée (à l'instar des chansons populaires). Dans le même esprit, pour la présentation de *The Lady* (Frank Borzage, 1925), histoire d'une danseuse à laquelle on arrache l'enfant qu'elle tentera toute sa vie de retrouver, un exploitant de l'Oregon propose d'utiliser une imposante copie de «The Mother» de J. A. Whistler (portrait d'une figure féminine mélancolique), et de faire chanter «Wonderful Mother»<sup>557</sup>. Selon la description faite dans cet article, le prologue du film de Borzage contient également, après une scène de pantomime dans une rue londonienne, une petite saynète dialoguée entre un policier et une dame. Ce dialogue s'organise autour d'une pointe finale qui fait office de slogan publicitaire, puisqu'il s'agit du titre du film.

Ces exemples montrent que les prologues conçus quelques années avant l'arrivée du parlant recouraient avec parcimonie à la parole, mais offraient des mises en scène complexes et variées de type «vaudeville». Ils ne correspondaient pas nécessairement à cette forme de présentation élémentaire qui deviendra par la suite le stéréotype du prologue où un personnage s'avance sur une scène pour s'adresser au public du film. Au contraire, les mises en scène conseillées par la First National excluent l'adresse au spectateur et mettent en place une diégèse «close», autonome par rapport à l'espace des spectateurs, annonçant ainsi le régime du film qui leur fera suite.

D'autres présentations attribuèrent par contre à la parole une fonction d'adresse. À l'occasion de la première du *Cabinet du Dr Caligari* (*Das Cabinett des Dr Caligari*, Robert Wiene, 1919) à New York le 3 avril 1921, l'exploitant de la prestigieuse salle du

557. «A Good Prologue for You Devised by Two Oregon Men», *Exhibitor's Herald*, Vol. XXII, N° 9, 22 août 1925, p. 40.

Capitol organisa un prologue et un épilogue scéniques. Un acteur qui se présenta comme le personnage auquel Francis raconte son histoire dans le film introduisit la première séquence, puis réapparut à la fin de la projection pour ajouter un *happy end*, affirmant que le jeune homme s'était totalement rétabli de sa folie<sup>558</sup>. Même si l'identité fictive de ce «locuteur» renvoie au monde du film, l'absence d'actions représentées sur scène fait de sa parole l'unique objet d'intérêt, et du public de la salle son seul destinataire. Cette composante caractéristique d'une parole-attraction est néanmoins contrebalancée par le contenu essentiellement narratif du texte. À l'instar de la stratification de l'énonciation filmique dans *Caligari*, cet encadrement *live* participe d'une logique de clarification des données du récit, l'intrigue étant plus ambiguë que ne le laisse entendre l'interprétation rationnelle et réaliste de ce prologue<sup>559</sup>.

En dépit de la fonction narrative du texte, la présence du locuteur sur la scène reste indéfectiblement liée à la voix-attraction, à l'inscription de la source vocale dans le même espace que les spectateurs. En termes de proxémique, l'attitude qui consiste à s'avancer puis à déclamer le texte instaure un rapport particulier au public qui se trouve implicitement désigné par le déplacement du locuteur dans sa direction. Par exemple, dans un prologue aux *Dix Commandements* (Cecil B. DeMille, 1923) produit par un promoteur anglais, Moïse descend de la montagne avant de venir délivrer son message au public<sup>560</sup>. Ce rapprochement spatial avec l'auditoire trouve un écho dans le texte déclamé: progressivement, le prophète s'abstrait du contexte de l'ancienne Égypte figurée par le Mont Sinaï pour s'adresser au «monde moderne» qui a connu «la foudre de la Guerre Mondiale», puis termine son discours, débuté au «je», en incluant le spectateur dans des tournures au pluriel («*nous* ne pouvons vivre sans la Loi», «*notre* histoire») et en établissant un pont entre l'époque actuelle (il utilise l'expression déictique «aujourd'hui») et celle de la transmission du Décalogue. L'auteur du descriptif de ce prologue précise qu'à la fin de son discours, Moïse regagne le mont et disparaît. Fragment délimité inscrit dans une succession d'attractions (parmi lesquelles on trouve

558. Selon David Robinson, *Das Cabinet des Dr Caligari*, London: BFI, 1997, pp. 47-48.

559. Voir Alain Boillat, «Statut énonciatif et fonction structurelle de l'enchaînement narratif dans *Caligari*», in *Limina/Le Soglie del film/Film's threshold*, op. cit., pp. 335-346.

560. H. Scott, «English Showman Joins Circle; Describes Presentation», *Exhibitor's Herald*, Vol. XXI, N° 12, 6 juin 1925, p. 51.

les manifestations du divin à grand renfort d'effets spéciaux), le moment de l'énonciation verbale s'affiche comme parole donnée au public, et cela d'autant plus fortement qu'il s'agit ici de la «bonne» parole. Tel est aussi le cas d'un autre prologue mentionné dans le même article. Prévu pour accompagner *L'Aigle des mers* (*The Sea Hawk*, Frank Lloyd, 1924), il prend pour cadre un bateau: le commandant s'avance, résume les événements, puis regagne l'équipage. L'individualisation du locuteur grâce à la mise en scène, accentuée dans un spectacle qui ne recourt par ailleurs que minoritairement à la parole, fonde la condition d'existence de la voix-attraction; le contenu des paroles relève quant à lui du narratif.

Ces quelques exemples de prologues du «muet» démontrent que le langage parlé, s'il n'y était pas prédominant, trouvait néanmoins à s'insérer dans les programmes et à côtoyer les films. L'univers sonore de ces spectacles périphériques constituait en lui-même une attraction, voire s'intégrait au film lorsque les chants *live* faisaient office de «ponts sonores» en se poursuivant sur le générique d'ouverture de l'écran<sup>561</sup>. En dépit de son déclin, la pratique des prologues vivants s'est poursuivie durant les premières années du parlant, conduisant ainsi à une juxtaposition de deux régimes vocaux parents mais fondamentalement différents en termes d'intégration de premier niveau.

L'intrigue du film *Footlight Parade* (Lloyd Bacon, 1933), significativement distribué en France sous le titre *Prologues*, est située dans le contexte des projections parlantes accompagnées de préludes scéniques. Ce film réalisé durant l'interrègne opère en quelque sorte une mise en abyme de l'intermédialité à l'œuvre dans les productions cinématographiques du début des années 30. Le premier plan de *Footlight Parade* nous montre un texte défilant sur la devanture d'une salle de cinéma, informant les passants grâce à ce dispositif électrique (plaçant ceux qui s'expriment par ce biais du côté de la «mécanicité») que «les producteurs [...] ne tourneront plus de films parlants; c'est la fin du muet». Un raccord-regard introduit ensuite le personnage principal interprété par James Cagney, Kent, un réalisateur de spectacles de music-hall. Le récit du film part donc du constat que la généralisation du parlant contrecarre le désir d'un individu de poursuivre sa pratique de la scène. Kent assiste à une projection précédée d'un prologue qui

561. R. Canjels, *art. cit.*, p. 311.

fait office de «bande-annonce» pour un film montré dans une autre salle. Le propriétaire lui confie que ce type de spectacle, trop cher, est devenu superflu à l'ère du parlant. Kent décide toutefois d'industrialiser les prologues vivants. Il devient alors un promoteur bouillonnant et inventif, dont les chorégraphies comprennent fréquemment, ainsi que nous le montrent les séquences de répétitions, des dialogues chantés. Le dernier quart du film de Bacon est composé d'une succession de trois prologues, dont la visualisation *in extenso* est motivée par un enjeu narratif: l'équipe de Kent doit faire ses preuves pour décrocher un contrat important avec un financier qui assiste aux performances de la troupe dans trois salles différentes. Ces séquences chorégraphiées ont en fait été réalisées par Busby Berkeley, dont on reconnaît l'esthétique basée sur l'éclatement de l'espace diégétique: les décors gigantesques dans lesquels évoluent les danseuses ne sont absolument pas vraisemblables en regard de l'espace scénique postulé par la diégèse. Le caractère ubiquitaire de la caméra – à laquelle il est impossible d'assigner l'œil d'un spectateur présent dans la salle – atteint son paroxysme dans les vues sous-marines. En faisant oublier de la sorte l'origine scénique du spectacle, ces prologues assurent en fait la liaison entre le vaudeville et la comédie musicale. Le principal point commun entre ces deux pratiques spectaculaires est bien sûr la danse, c'est-à-dire le mouvement; le dernier prologue du film se termine significativement sur le gros plan d'un *flip book*. La voix-attraction occupe néanmoins une place importante dans l'un comme dans l'autre. Le contexte dépeint dans le film de Bacon et Berkeley illustre l'hétérogénéité de l'interrègne. Lorsque l'oralité cède un peu plus la place à l'écrit, nous trouvons alors l'un des ancêtres de la voix-over inaugurale: le prologue filmé.

## **V.5 LES PROLOGUES DES PREMIERS PARLANTS: LA SCÈNE A L'ECRAN**

Les prologues *live* trouveront un équivalent sous une forme fixée dans les courts métrages réalisés dans les premières années du parlant. Il en va par exemple ainsi du célèbre discours dans lequel William Hays annonce l'avènement du parlant, qui, lors de la présentation publique officielle du Vitaphone du 6 août 1926 à New York, constitua un prélude aux *Vitaphone Shorts*, courts métrages dominés par une esthétique attractionnelle. Cette allocution fit



d'ailleurs doublement office de prologue puisqu'elle se situait en première position au sein d'une série de courts métrages musicaux ou chantés qui, dans l'ensemble, ne formaient eux-mêmes qu'un avant-programme dans une séance principalement consacrée au long métrage *Don Juan* d'Alan Crosland (projeté avec musique et quelques rares bruitages au synchronisme large, mais sans paroles).

Dans la bande parlante montrant William Hays, tous les paramètres visuels et sonores concourent à associer la voix à un geste inaugural de *présentation*. Mise en scène et mise en cadre participent pleinement de la référence à une situation d'énonciation particulière, celle de la *conférence*: le locuteur est initialement situé entre un pupitre et le sigle de la Vitaphone Corporation, éléments qui soulignent la position d'autorité qui lui est conférée. Hays est présenté (et se présente) comme un patron s'adressant à un hypothétique conseil d'administration, alors même que son discours est rendu public au-delà des murs de la Warner. En imitant l'intimité de la communication interne à l'entreprise, Hays thématise la fonction de divulgation à grande échelle du savoir, démarche « performative » s'il en est, puisque son allocution vise précisément à convaincre l'auditoire qu'il s'agit là de la vocation du système Vitaphone. En dépit de ce cadre spécifique que connote un décor rudimentaire, l'usage liminaire de la voix n'est pas sans évoquer les (plutôt rares) prologues des spectacles vivants qui recouraient à la voix parlée. Par exemple, Hays se retire à l'issue de son discours au fond de la pièce après une légère révérence, comme lorsque les *showmen* des prologues scéniques laissent la place à la projection cinématographique. Il n'est pas anodin que cette profération que l'on considère mythiquement comme la « première » parole synchrone de l'histoire du cinéma soit une adresse, élément caractéristique du « cinéma des attractions ». La volonté de créer une « illusion de la présence » que Kessler et Lenk observent dans les films de Méliès me semble également sous-tendre le discours de Hays, qui peut être conçu comme une performance<sup>562</sup>. Cette dernière semble aujourd'hui fort banale, tant ce type d'adresse au public s'est généralisé avec la représentation télévisuelle des speakers. La comparaison avec la transmission cathodique n'est d'ailleurs pas fortuite: comme l'a relevé Michel Chion, le discours inaugural de

562. Cf. Frank Kessler et Sabine Lenk, « L'adresse-Méliès », in *Georges Méliès, l'illusionniste fin de siècle?*, op. cit., p. 197.

Hays, « bien qu'enregistré, joue la carte de la « retransmission » du temps réel, de la simulation d'une présence en direct »<sup>563</sup>. Ce paradoxe, qui reflète celui de la présence-absence inhérent à la technologie phonographique, n'est pas tant une anticipation du *broadcasting* que l'indice d'une volonté du locuteur d'affirmer sa présence (en tant que « corps parlant » interagissant avec ses auditeurs) par rapport à la nature mécanique de l'enregistrement. En démontrant le primat de la parole humaine, Hays promeut le système sonore qu'il décrit. Or l'émission de cette voix synchrone s'accompagne d'une configuration spatiale qui renvoie elle-même à l'un des contextes dans lesquels la voix vive prend place. Bien qu'il évite les regards marqués à la caméra, Hays effectue d'amples mouvements de la tête et des bras, comme s'il embrassait du regard un public situé autour de lui. Cette gestuelle contribue à faire oublier la technologie qu'il présente au profit de la performance orale, la référence à un contexte scénique favorisant l'occultation du tournage et de la reproduction phonographique.

Une telle simulation de pratiques *live* apparaît dans la partie introductive de plusieurs *talkies*. Situé dans les marges du film, le couple prologue/épilogue s'avère particulièrement apte à se détacher de l'univers diégétique pour faire apparaître un locuteur qui se tourne vers le spectateur et présente des informations sur le film « en tant que film ». C'est donc naturellement aux seuils filmiques que l'on trouvera un type communicationnel régi par les dominantes phatique et métadiscursive qui caractérisent la voix-attribution. La pratique du prologue filmique s'enracinant dans les spectacles de la scène, il n'est pas rare qu'elle soit marquée par une forte théâtralité. Cet « intertexte » constitue par conséquent une base référentielle commune qui favorise la « prise de contact » et renforce l'effet de proximité. Sans postuler une influence sur l'ensemble de la production cinématographique, j'aimerais évoquer quelques films américains qui témoignent d'une résurgence de la pratique du prologue, courante dans les spectacles de vaudeville et dans les prologues *live* du muet.

Avant d'en venir à des productions du début des années 30, il me paraît intéressant d'évoquer le cas d'un film qui se présente ostensiblement comme l'intégration (primaire et secondaire) d'un « acte » parlé sous forme de prologue : il s'agit du *Masque de fer* (*The*

563. M. Chion, *Le Son au cinéma*, op. cit., p. 65.

*Iron Mask*, Allan Dwan, 1929). Réalisé dans une année-charnière, le film de Dwan est symptomatique de l'hybridation des premières projections parlantes du point de vue de leur régime spectaculaire et vocal. Ce film est un cas particulier dans la mesure où, tourné en muet en 1929, il est sonorisé dans un second temps selon une méthode qui n'a pas consisté à établir, comme cela se faisait fréquemment à l'époque, une version parlante qui aurait existé parallèlement à la version muette, mais uniquement à ajouter un prologue et un intermède parlants. La synchronisation vocale s'applique donc exclusivement à deux « attractions » périphériques au récit filmique. Dans le prologue du *Masque de fer*, les deux réceptions de « l'attraction » (le « son comme attraction » et l'attraction du spectacle de vaudeville) se confondent. Pour cette apparition ponctuelle, la parole n'est pas prêtée à un *showman* extérieur au film, mais elle est placée dans la bouche du personnage de d'Artagnan. L'attraction de cette séquence inaugurale réside, comme dans les *Vitaphone Shorts* où s'exhibaient les célébrités du vaudeville et de l'opéra, dans la présence même de la star de l'écran (Douglas Fairbanks) qui s'exprime en voix synchrone. Dans l'entretien qu'il accorde à Peter Bogdanovich, le cinéaste relate ainsi cet aspect de la genèse du film :

« C'est à cette époque que le parlant est apparu, et j'ai dit à Doug (Douglas Fairbanks) : « Tu sais, l'arrivée du parlant est notre épée de Damoclès, il va falloir les prendre de vitesse. » Nous avons donc tourné un prologue dans lequel Doug en d'Artagnan déclamaient une tirade poétique. Puis le film se déroulait, et à la fin on a mis un épilogue dans lequel le vieux d'Artagnan faisait ses adieux. Cela donnait au film une dimension formidable, tout à fait moderne. »<sup>564</sup>

Il va de soi qu'une telle version des faits, donnée quarante ans après la réalisation du film, s'accompagne d'une certaine idéalisation : on y retrouve par exemple l'affirmation d'une prescience de l'hégémonie du parlant qui est caractéristique des récits rétrospectifs portant sur cette période. Le fait que Dwan mentionne le couple prologue/épilogue est révélateur de la manière dont l'insertion de

564. Peter Bogdanovich, *The Last Pioneer*, New York : Praeger, 1971, repris et traduit in Giorgio Gosetti, *Allan Dwan. La Légende de l'homme aux mille films*, Paris ; Locarno : Éditions des Cahiers du Cinéma ; Festival International de Locarno, 2002, p. 65.

scènes parlées a été pensée, c'est-à-dire en tant que marges du film. Toutefois, cet encadrement ne correspond pas au montage final (du moins dans la version disponible aujourd'hui), «l'épilogue» se situant en fait à la fin de la première partie (à 56 minutes du début du film), après l'intertitre «Vingt ans après». Ce second moment en voix synchrone introduit de la sorte une phase à laquelle l'importante ellipse qui précède et la référence à l'ouvrage d'Alexandre Dumas (*Vingt ans après* étant la suite des *Trois Mousquetaires*) confèrent le statut de «nouveau début», ou d'écho (la devise «Tous pour un, un pour tous» est reprise). La voix y acquiert donc bien une fonction démarcative (on l'entend après le générique, puis après la première partie) et présentationnelle qui rappelle passablement les prologues scéniques. Il est donc surprenant que le cinéaste qualifie cette pratique de «tout à fait moderne», alors qu'il utilise simultanément une terminologie empruntée au domaine du théâtre («déclamer», «tirade») pour qualifier le speech de Fairbanks. Il faut comprendre que si «modernité» il y avait, elle résidait plutôt dans l'utilisation même du son synchrone que le public pouvait s'attendre à trouver au moment de la sortie du film – en février 1929, alors que le film de Dwan était en cours de réalisation, Valentin Mandelstamm annonçait avec optimisme que «dans *Le Masque de fer* [...] les acteurs parleront» – la «dimension formidable» étant alors conférée par le statut d'«attraction» de la production de cette parole par des moyens mécaniques<sup>565</sup>. Il est utile de décrire quelques caractéristiques de ce prologue somme toute fort proche des «Film parlants» produits quinze ans auparavant par Gaumont, car on peut ainsi saisir combien ce type d'ajout sonore témoigne de l'inscription des premières bandes parlantes d'après 1926 dans la filiation de certaines manifestations scéniques de la voix-attraction. D'une durée d'une minute et vingt-cinq secondes, ce prologue est composé d'un unique plan fixe montrant frontalement une scène dont l'arrière-plan est occupé par quatre personnages qui se tiennent debout, le visage face au «quatrième mur», devant un fond représentant en aplat une petite ville. Le décor scénique, délimité par des rideaux, des fioritures et un ruban sur lequel figure le titre «The Iron Mask», redouble exactement le cadre filmique: l'espace profilmique semble

565. Valentin Mandelstamm, «L'Avènement des Films parlants et synchronisés», *Cinéma-gazette*, N° 5, 1<sup>er</sup> février 1929, p. 197.

se donner à voir comme l'intégralité de l'espace diégétique. Douglas Fairbanks s'avance, sort son épée du fourreau d'un mouvement sec, lance son chapeau hors-champ (ou plutôt dans les coulisses) et commence son allocution, alors qu'une musique de fond apparaît discrètement, s'affirmant à la fin du prologue pour constituer le finale. Notons que durant tout le passage, les trois autres personnages restent quasi immobiles (si ce n'est pour obéir à l'appel de d'Artagnan à lever l'épée) et muets : ils sont relégués à la fonction de figurants, d'éléments du décor, alors que le locuteur occupe une position privilégiée puisqu'il est situé à l'avant-plan, exactement au centre de l'image. Le rituel de la présentation se déroule donc comme certains prologues *live* que j'ai mentionnés ci-dessus. Le *showman* descend du sommet de la pièce d'artillerie sur laquelle il « prenait la pose » – se désolidarisant ainsi du « tableau » muet qu'il contribuait à composer à l'ouverture du plan – s'avance, fixe son public et s'exprime. Cette mise en scène me paraît jouer du contraste entre mutité et paroles. Dans la description qu'il fait de ce prologue et de sa genèse, Maurice Leloir, un historien français engagé comme conseiller historique sur le film, évoque l'apparence volontairement surannée du décor (contrastant avec la déclaration du cinéaste qui considère ce prologue comme « moderne ») et le caractère totalement figé de l'image avant que ne surgisse la parole :

« Ce prologue est présenté comme une tapisserie ancienne avec sa bordure ornementale et un sujet central sous le titre « The Iron Mask ». Dans une redoute fortifiée, les quatre mousquetaires sont représentés faisant serment d'aide mutuelle et croisant leurs épées. Cette tapisserie est une aquarelle que j'exécutai. Au bout d'un moment le sujet du centre prend vie et, toujours encadrés par la bordure, les personnages s'animent, deviennent vivants. D'Artagnan faisant deux pas en avant dit un prologue en vers que le Vitaphone fait entendre. »<sup>566</sup>

Le prologue du *Masque de fer* est sous-tendu, comme l'explique Leloir, par l'idée de donner la vie – c'est-à-dire non seulement le mouvement, mais aussi la parole – à une image fixe et muette.

<sup>566</sup> Maurice Leloir, *Cinq mois à Hollywood avec Douglas Fairbanks*, Paris : J. Peyronnet Éditeurs, 1929, p. 87.

Le décor se réfère à des modes de représentation antérieurs au cinéma: d'abord les images fixes (l'aquarelle, la tapisserie) puis, avec les évolutions de Fairbanks dans l'espace de la scène, le théâtre. L'emprunt à d'autres formes d'expression était en effet courant chez les promoteurs de prologues à l'époque du muet. La continuité dont témoigne le film de Dwan avec une telle pratique scénique contribue à renforcer la théâtralité qu'induisent les pratiques filmiques « primitives » que ce prologue met en œuvre. Tout concourt, dans ce prologue, à mettre en valeur la profération de parole, car l'attention des spectateurs est entièrement focalisée sur le personnage joué par Fairbanks. La précision de la synchronisation, d'abord posée au niveau des bruits (l'épée fendant l'air à trois reprises), est ensuite confirmée avec le début de la déclamation. On peut imaginer que cet élément technique devait à lui seul faire office d'« attraction » à la sortie du film, ceci d'autant plus que l'émergence de la parole s'opérait souvent, dans les films parlants de cette époque, quelque temps après le début du film (par exemple dans *Chantage*, réalisé la même année par Alfred Hitchcock, qui joue sur le « suspense » de l'apparition vocale). Par ailleurs, la diction aux accents poétiques (intonation emphatique, accentuation de certaines syllabes, respirations en fin de « vers », etc.) introduit une composante réflexive sur le plan verbal qui renforce la valeur attractionnelle de la voix. La gestuelle de Fairbanks, essentiellement réduite aux mouvements du bras droit – puisqu'il est campé dans la position rigide de celui qui détient l'autorité de la parole, se met également au diapason de la dominante attractionnelle et du regard-caméra. En effet, l'acteur, tourné vers le public, ne fait aucun geste à l'attention des autres personnages, mais souligne les pronoms d'adresse du texte par de larges mouvements du bras en direction d'un auditoire virtuel (« J'en appelle à votre patience... », « Suivez-moi dans la France de jadis », etc.). Ensuite, en point d'orgue à sa déclamation, il lève l'épée en répétant « Come on », ainsi que le ferait un chanteur qui voudrait enjoindre son public à continuer de chanter après la fin de sa prestation. Cette incitation semble également suggérée par la musique qui, amplifiée, se prolonge au-delà du fondu au noir. Ce prologue est bien un lieu de « voix-attraction », tant du point de vue des caractéristiques vocales, de l'exhibition de la technologie que des modalités de l'adresse. Le film de Dwan semble être en outre attaché à un destin singulier quant à l'utilisation de la voix: en 1952,

Lippert Pictures distribua une version écourtée et sonorisée qui était accompagnée d'un commentaire écrit par Richard Llewelyn et dit *over* par le fils de l'acteur principal, Douglas Fairbanks Jr. Ainsi la voix-attraction de la version originale a fait place à une voix-narration (comme en témoigne le bref extrait offert en supplément du DVD édité par Kino On Video, 2002), signe de l'adaptation du régime vocal original à la période classique. On verra avec l'exemple suivant que ce mouvement vers l'intégration narrative s'opère déjà au milieu des années 30.

On trouve également l'intervention liminaire d'un *showman* dans les deux célèbres «films de monstre» que la Universal produisit sous l'égide de Carl Laemmle en 1931 : *Frankenstein* de James Whale et *Dracula* de Tod Browning. Comme il s'agit d'adaptations de textes qui ont également donné lieu à des représentations sur les scènes de théâtres populaires, ces films se prêtent à une comparaison avec ce qui a été dit du spectacle de type «vaudeville». Par ailleurs, la suite du film de Whale réalisée quatre ans plus tard, c'est-à-dire à la fin de l'interrègne (*La Fiancée de Frankenstein/The Bride of Frankenstein*, 1935), présente en son prologue un régime vocal totalement opposé à celui du film de 1931 : le passage d'un prologue à l'autre me permettra d'illustrer les deux pôles voix-attraction/voix-narration, tout en soulignant leur ancrage historique.

Il semble bien que durant les premières années du parlant, les films d'horreur aient connu un usage particulier de la voix qui s'explique peut-être par la relation paradoxale que ce genre entretient avec le public (attirance *vs.* répulsion, confort *vs.* malaise, identification *vs.* agression, etc.) : dans ce contexte ambivalent, la parole intervient simultanément sous la forme d'une mise en garde et d'une menace. Ainsi, le second film entièrement parlant produit par la Warner, *The Terror* (Roy del Ruth, 1928), comprend un générique parlé que récite, en voix synchrone, un personnage vêtu d'une cape et d'un masque<sup>567</sup>. On constate là une parenté entre prologue et générique, et la propension que connaissent les films horribles à introduire immédiatement le public dans un univers étrange. Autre exemple : avant même que le générique de *Frankenstein* (James Whale, 1931) ait commencé, un présentateur sort de derrière le rideau d'une scène filmée frontalement pour s'adresser

567. Selon Jean-Pierre Coursodon, *La Warner Bros*, Paris : Centre Georges Pompidou, 1991, p. 85.

au public. Vêtu d'un costume (nœud papillon, fleur à la boutonnière et mouchoir blanc à la poche), s'adressant posément et avec affabilité à son auditoire virtuel, ce présentateur, interprété par l'acteur Edward Van Sloan qui incarne également un personnage dans le film (le Dr Waldman), a tout du *showman* tel qu'on peut en trouver dans les spectacles respectables de Broadway ou dans les prestigieux théâtres de vaudeville à partir des années 1890. L'absence d'autres centres d'intérêt dans l'image porte inévitablement toute l'attention du spectateur sur ce locuteur et sur ses paroles. Une fois ce présentateur «entré en scène» – car il s'agit là d'une entame toute théâtrale – la caméra s'avance progressivement jusqu'à le cadrer à la taille. Cette plus grande proximité avec l'acteur nous permet de constater que le regard de Van Sloan ne croise jamais l'axe de la caméra, mais mime les mouvements oculaires d'un conférencier qui jetterait sans cesse des regards à droite et à gauche pour couvrir toute une salle. On ressent une stratégie d'évitement du regard-caméra, celui-ci étant reporté sur un public diégétique virtuel, place que le présentateur assigne indirectement au spectateur du film.

Cette question du regard mérite d'être discutée. Plutôt que de parler d'un «regard-caméra», je propose de qualifier ce type d'adresse «*regard-public*»: il s'agit en effet d'un regard dirigé vers les spectateurs (que l'on peut difficilement qualifier de «diégétiques», tant la diégèse filmique est encore sommairement établie) supposément situés devant le locuteur et auxquels certaines paroles se réfèrent. Certes, le regard-public confine parfois momentanément au regard-caméra et provoque, à l'instar de ce dernier, un effet d'interpellation du spectateur du film. On distinguera néanmoins un regard-caméra délibéré qui rompt l'illusion de la représentation par l'exhibition du pôle machinique d'un regard-public qui, au contraire, participe du simulacre audiovisuel en faisant croire au spectateur du film qu'il appartient à un espace situé devant une scène, et qui procède par conséquent d'une simulation du pôle humain du dispositif. Jean Châteauevert a proposé une distinction entre adresses préfilmique (audience située dans l'espace du tournage) et postfilmique (audience de la salle de cinéma) qui rend compte des deux types que j'évoque ici. Il mentionne par ailleurs un second niveau de différenciation qui me paraît important: dans le cas d'un regard d'adresse préfilmique, l'adresse *individuelle* vise un auditoire précis car visualisé, alors que l'adresse *plurielle*



concerne un ensemble de personnes virtuelles. Or l'écart qui sépare ces deux situations est peu pris en compte par Châteauevert, alors qu'à mon sens, la frontière devient fort ténue entre l'adresse préfilmique et l'adresse postfilmique du moment qu'il y a «virtualisation» d'un public diégétique rejeté hors-champ. Il s'agit plutôt d'un cas intermédiaire («regard-caméra au public») où le spectateur postulé par l'adresse préfilmique, c'est-à-dire celui qui est «présent au lieu et au temps du spectacle filmé»<sup>568</sup>, se voit fictivement assimilé au spectateur du film. Il s'agit alors, si on interprète cette situation dans le contexte théorique qui est le nôtre, d'une stratégie pseudo-attractionnelle visant à simuler les conditions de réception des spectacles vivants.

Le regard-public d'Edward Van Sloan dans le prologue de *Frankenstein* trouve un écho dans le texte qu'il profère. Ce dernier débute par une interrogation purement phatique («Comment allez-vous?»), puis enchaîne sur une introduction réflexive: «M. Carl Laemmle estime qu'il serait désobligeant de présenter ce film sans un aimable avertissement.» Cette phrase institue le présentateur en porte-parole du producteur: le prologue qui était conçu à l'époque «muette» par l'exploitant se met ici explicitement au service de l'instance de production du film. Le *showman* qualifie son propre speech d'«avertissement»: il ne s'agit bien sûr pas de faire fuir le public de la salle, mais au contraire d'éveiller sa curiosité tout en fixant les termes du «contrat de lecture». En avertissant le public du caractère horrifique de ce qui va suivre, le présentateur conclut une sorte de promesse. Ainsi, si une phrase synthétique annonce vaguement le contenu narratif du film («Frankenstein, un homme de science qui voulut créer un homme à son image»), il s'agit avant tout de préparer le public à l'atmosphère du film, fonction qui importait déjà aux bonimenteurs du cinéma des premiers temps et aux acteurs de prologues *live*. Cette parole inaugurale vise donc plus à toucher le public au niveau de l'affect qu'à fournir des informants narratifs. Selon un principe de surenchère que l'on trouve également dans les intertitres du récit-cadre du *Cabinet du Dr Caligari*, le présentateur annonce l'effet que ne manquera pas de produire le film («Je pense qu'il vous fera peur. Il pourrait vous choquer. Il se peut même qu'il vous terrifie»). Certes, les précautions formulées dans ce prologue ont également

568. J. Châteauevert, «Regards à la caméra...», *art. cit.*, p. 222.

une origine plus pragmatique : d'après l'historien Rudy Behlmer, la production entendait prévenir ainsi les objections de groupes religieux (voir les commentaires de Behlmer dans les suppléments de l'édition DVD de *Frankenstein*, Universal, 2002). Il transparaît néanmoins des phrases de ce prologue (moult fois réécrit, notamment par John Huston) que les intentions de ses concepteurs ne se bornent pas à parer une éventuelle accusation de présomption divine, mais aussi à prédisposer le public à l'attraction, à l'instar des aboyeurs de Luna Park qui « chauffent » le public déjà assis dans le wagon d'un « train fantôme ». Dans *Frankenstein*, le but n'est pas d'exhiber la qualité de l'invention technologique, mais d'utiliser un mode de représentation popularisé par les spectacles vivants pour mieux accrocher le public et instaurer une proximité qui crédibilise l'action à venir. Une fois cet avertissement transmis, le présentateur en smoking quitte le champ, laissant la place au générique après un bref plan noir – signe du passage à « l'acte » suivant – sur lequel débute déjà la musique.

Il est étonnant de constater que le prologue de *Frankenstein* trouve un « petit frère » dans l'épilogue de *Dracula* (Tod Browning, 1931), film réalisé la même année en multiversions par la même maison de production et avec le même présentateur, Van Sloan. À nouveau, l'acteur interprète un personnage du film (Van Helsing, qu'il incarnait déjà sur les planches Broadway dans l'adaptation du roman de Bram Stoker) que le spectateur est ici en mesure d'associer au locuteur puisque l'adresse, semblablement filmée devant un rideau de théâtre, fait suite au film. David Skal, co-auteur d'une monographie de Browning – on y apprend notamment combien ce dernier s'est familiarisé avec les spectacles vivants populaires avant de devenir cinéaste (il fut aboyeur, chanteur de vaudeville, acteur de *blackface minstrels*, ...) <sup>569</sup> – livre dans l'édition DVD de *Dracula* (Universal, 2002) le commentaire suivant à propos de cet épilogue qui a été retiré des copies du film à partir de sa ressortie en 1938 <sup>570</sup> :

« À la fin de chaque représentation [de l'adaptation théâtrale], au moment où le rideau tombait, le professeur Van Helsing s'avançait et interrompait les applaudissements en levant les bras

569. Skal et Savada, *Dark Carnival. The Secret World of Tod Browning*, New York ; London : Anchor Books, 1995, pp. 28 et 31.

570. *Ibid.*, p. 151.

pour dire quelques mots. Un des membres de l'équipe, Ivan Butler, me raconta qu'en province le public, qui connaissait le discours par cœur pour avoir assisté aux différentes tournées, le récitait à voix haute, comme pour la version 1920 du *Rocky Horror Picture Show*. Ce discours fut inclus dans le film original mais censuré par la suite, de peur qu'il n'offensât des groupes religieux en promouvant le fantastique.»

La genèse de cet épilogue filmique est explicitement ancrée dans une pratique scénique, et associée au contexte du spectacle populaire où la voix vive acquiert le pouvoir de gagner les spectateurs. Le regard-public et la voix d'adresse s'originent dans ce type de représentation. À l'instar de nombreuses voix-over de films d'horreur ultérieurs, l'épilogue verbal du film de Browning crée un effet de relance, maintient l'horreur dans cet espace hors-film afin que le spectateur en reste marqué après avoir quitté la salle. À l'argument avancé par Skal pour expliquer la suppression de cet épilogue, on peut ajouter une motivation liée au nouveau contexte qu'a connu la ressortie de *Dracula*: en 1938, alors que la voix-over est sur le point de connaître un premier essor, ce type d'adresse ne convient plus aux pratiques qui se sont instaurées de façon dominante à la fin de l'interrègne. Les discours en voix synchrone de *Frankenstein* ou de *Dracula* contribuent peu à la mise en place ou à la clôture du récit: plus réflexifs et phatiques que référentiels, ils agissent ostensiblement en tant que facteurs de réglage de la lecture des spectateurs. L'intégration totale du verbal inaugural au discours filmique et sa soumission complète aux impératifs narratifs s'effectuera avec les ouvertures de films dans lesquelles la voix abandonne tout ancrage dans l'espace scénique pour devenir *over*, à l'exemple de la suite de *Frankenstein*.

#### **V.6 DU MONSTRE A SA FIANCEE:**

##### **«CREATION» FRANKENSTEINIENNE DE LA VOIX-NARRATION**

La voix-attraction endosse souvent une fonction inaugurale: on la trouve surtout dans les prologues et les génériques, voire en dehors des films. L'hégémonie du narratif a progressivement conduit, durant l'interrègne, au remplacement de la voix-attraction initiale par une voix-narration sous forme *over*. Dans le répertoire chronologique (non exhaustif) établi par Sarah Kozloff des

productions américaines comportant une voix-over, *La Fiancée de Frankenstein* de 1935 est le troisième film cité. Il me paraît intéressant de réfléchir sur les liens que l'une des plus anciennes voix-over du cinéma hollywoodien entretient avec l'adresse en voix synchrone qui a marqué les premières bandes parlantes. Dans notre démonstration, ce film fera office de pivot entre la voix-attraction étudiée dans le vaudeville et la voix-narration dont il sera question dans la suite de l'ouvrage.

La comparaison du *Frankenstein* de 1931 avec son *sequel* est d'autant plus aisée que la voix-over de ce dernier n'est ni continue, ni récurrente, mais cantonnée à la partie initiale du film. En fait, le prologue de *La Fiancée de Frankenstein* – situé après les cartons du générique, ce qui signale déjà sa pleine participation à la mise en place de la diégèse – comporte deux voix-over, l'une masculine (celle du personnage de Lord Byron), l'autre féminine (celle de Mary Shelley). La première, interprétée par Elsa Lanchester qui incarnera le personnage-titre, se contente de *rappeler* l'épisode précédent afin d'en *appeler* la suite, alors que la seconde *donne naissance* au récit. Cet « accouchement » (souhaité par la figure masculine mais réalisé par le personnage de la romancière) prend un sens tout particulier dans ce film qui place son récit explicitement sous l'autorité énonciative d'un personnage féminin (en tant qu'instance déléguée), redonnant à la femme un pouvoir qu'elle ne possédait pas dans l'histoire du premier *Frankenstein*. L'accès à la parole d'une part, la mise en place d'un récit enchâssé d'autre part confèrent une position « focale » au personnage de Mary Shelley. À cet égard, Skal a souligné que le film de Whale propose une « lecture » du roman de Mary Shelley qui en actualise le sous-texte féministe lié à l'anxiété d'une autrice envers le désir du savant de trouver une nouvelle méthode de procréation qui ne nécessite pas le principe féminin<sup>571</sup>. *La Fiancée de Frankenstein* opérerait donc une restauration du « point de vue auctorial » de la femme dont Amy Lawrence a montré que les personnages féminins de la plupart des productions hollywoodiennes étaient privés<sup>572</sup>. Il faut cependant noter que l'emprise du discours de la narratrice sur le récit audiovisuelisé est limitée par les conventions du prologue :

571. David J. Skal, *The Monster Show. A Cultural History of Horror*, New York: Faber & Faber, 1993, p. 184.

572. Amy Lawrence, *Echo and Narcissus, Women's Voice in Classical Hollywood Cinema*, Berkeley; Los Angeles; Oxford: University of California Press, 1991.

plus jamais par la suite nous n'entendrons ni ne verrons l'actrice Elsa Lanchester dans le rôle de Mary Shelley. *In fine*, ce sont tout de même les tourments du docteur Frankenstein qui prennent le pas sur la «voix narrative» de la romancière.

Dans *La Fiancée de Frankenstein*, la référence au médium que l'on trouve dans les prologues de son *prequel* et de *Dracula* est déplacée du film sur sa source littéraire: Mary, l'auteur du roman *Frankenstein*, s'entretient avec Byron et son futur époux Percy Shelley de son œuvre d'épouvante encore non publiée, puis décide de leur raconter la suite de l'histoire. Ce déplacement n'est pas fortuit: le prologue, qui ne présente aucune parenté esthétique avec un espace scénique, élabore une diégèse filmique, celle de la situation-cadre. Un travelling avant au gré duquel la caméra, d'abord située à l'extérieur de la villa, pénètre dans le séjour par la fenêtre, constitue un signe manifeste de l'«entrée dans un monde», fonction que relaie ensuite un fondu enchaîné. Certes, les motivations extrafilmiques du prologue de *Frankenstein* subsistent, puisqu'on y retrouve la prévention envers d'éventuelles attaques de la part d'instances religieuses. Cette justification subit néanmoins le même transfert du film au livre, Mary Shelley déclarant que «les éditeurs n'ont pas compris que [son] histoire avait une morale». Contrairement au prologue du film de 1931, la séquence d'ouverture de *La Fiancée de Frankenstein* ne désigne pas le film en tant que film. L'existence matérielle du support filmique s'efface d'entrée de jeu, comme le fera ensuite le corps de la narratrice.

En dépit de la situation dialoguée, la voix-action est faiblement présente dans ce prologue: reclus dans un salon par une nuit d'orage, les personnages n'entreprennent rien d'autre que de converser, puis d'écouter un récit conté devant la cheminée. Byron s'exprime essentiellement sur le mode de l'évocation romantique avant de devenir le premier narrateur. Or, dès l'instant où il commence à relater les événements majeurs de *Frankenstein*, sa voix se fait *over* et les images, extraites du film de 1931, illustrent ses dires. Les images enchâssées ne sont pas muettes, même si le monteur a évité toutes les scènes dialoguées (on entend seulement des pleurs, le grognement de la bête, etc.). La musique de Waxman, déjà présente dans la situation-cadre, souligne les actions principales en même temps que la voix-over. *La Fiancée de Frankenstein* recourt pour ce collage d'extraits au mixage de différentes pistes-son, technique qui est apparue entre la réalisation de *Frankenstein* et celle de

sa suite<sup>573</sup>. Ainsi la voix-over peut-elle se détacher sur un fond « audiovisualisé » comme si elle nous guidait dans ce monde. Cette compilation des moments-clés du film, accompagnée d'un discours *over* élogieux (Byron vantant les mérites de l'œuvre de Shelley), n'est pas sans rappeler un autre type de production filmique encore plus « périphérique » que le prologue, mais similaire en termes de brièveté, d'autonomie et d'emplacement (première partie du programme) : la bande-annonce. Le montage rapide et elliptique du premier enchâssement de *La Fiancée de Frankenstein* nous donne l'impression d'assister à une présentation publicitaire du film de 1931. Bien qu'il s'agisse d'un rappel, l'effet d'annonce demeure : la valorisation qui est faite du récit de *Frankenstein* vaut également pour le film que nous nous apprêtons à voir, dont « l'auteur » est également Mary Shelley.

La comparaison de cet enchâssement liminaire avec la pratique de la bande-annonce nous permet de poursuivre la filiation intermédiaire que j'ai esquissée jusqu'ici puisque, ainsi que l'étudie Vinzenz Hediger dans son ouvrage consacré à l'histoire du *trailer* américain, ce type de promotion des films donna lieu à des pratiques mixtes qui alliaient projection et représentation scénique. Hediger illustre notamment ce spectacle hybride en mentionnant un exploitant du Massachusetts qui, en 1931, faisait intervenir sur scène un double de Dracula au milieu de la projection de la bande-annonce du film de Browning<sup>574</sup>. Pour « donner vie » à ces monstres improbables, leur présence effective dans la salle semblait être du meilleur effet. Hediger cite en outre un article du *Motion Picture World* de 1927 concernant la publicité pour un film produit par la MGM, *The Show*, dont le titre revendique l'influence d'un certain type de spectacle : « *Made a Vaudeville Act Trailer for The Show* », prône l'exploitant du State Theatre de Los Angeles<sup>575</sup>. C'est toutefois autour de 1928 que se généralise à grande échelle la publicité pour les films, la Warner consacrant cette année-là un département spécifique à cette activité. La standardisation de la fabrication des bandes-annonces est donc contemporaine de la transition au parlant, ce qui, au-delà des raisons économiques liées à la réorganisation des studios qui aboutit à la formation des cinq

573. Selon Barry Salt, *Film Style and Technology: History and Analysis*, London: Starword, 1983, p. 281.

574. Vinzenz Hediger, *Verführung zum Film*, Marburg: Schüren Verlag, 2001, p. 108.

575. *Ibid.*, p. 130.

*majors* au cours des années 20, n'est probablement pas sans liens avec le potentiel qu'est susceptible d'offrir la voix dans le cadre d'un discours de type persuasif. D'ailleurs, comme le montre Hediger, les bandes-annonces firent assez vite primer la narration. Dans l'un de ses articles de 1929 sur «l'avènement des films parlants et synchronisés», le Français Valentin Mandelstamm est d'ailleurs attentif à l'apparition conjointe du son et de cette pratique promotionnelle lorsqu'il note qu'«une autre innovation, très en vogue maintenant pour la distribution et la publicité des bandes, consiste dans les *Trailers*». La description qu'il fait de ce type de bandes souligne combien elles poursuivent l'esthétique des présentations orales vivantes :

«[Ce] sont des films de 30 à 50 mètres dans lesquels, selon les circonstances, les stars, les metteurs en scène et différents artistes ou techniciens qui ont contribué à l'exécution de l'ouvrage apparaissent à l'écran, avant la présentation, *s'adressent directement à l'auditoire*, racontent leurs impressions, [...] créent ainsi, dès l'abord, une *atmosphère cordiale et «réceptive»*.<sup>576</sup>

Pour Mandelstamm, les *trailers* usent donc de la parole pour rétablir à l'écran la *présence* des performances scéniques. La simultanéité du développement du parlant et des bandes-annonces n'est pas fortuite : tous deux procèdent de la même logique de réactualisation des caractéristiques du spectacle vivant. Ainsi *Le Chanteur de jazz* a-t-il droit à un *trailer* sous la forme traditionnelle d'un présentateur en smoking situé devant les rideaux d'une scène de théâtre. D'après la description qu'en fait Hediger, ce *trailer* ne s'en tient toutefois pas à la présentation scénique que l'on trouvera dans le prologue de *Frankenstein*. En effet, des plans sont insérés qui renvoient d'une part à la première du film à laquelle assistent les stars, d'autre part à quelques extraits chantés qui placent ostensiblement le spectateur de la bande-annonce dans la position d'un spectateur diégétique<sup>577</sup>. Ce second principe annonce la construction spatiale propre à l'intégration des moments de chant dans le film de Crosland, puis dans certaines comédies musicales.

576. V. Mandelstamm, «L'Avènement des Films parlants et synchronisés», *Cinémagazine*, N° 5, 1<sup>er</sup> février 1929, p. 282 (je souligne).

577. V. Hediger, *Verführung zum Film*, *op. cit.*, pp. 110-111.

La réflexivité dont font preuve les bandes-annonces auxquelles se réfère Mandelstamm en 1929 – que l'on retrouvera quelques années plus tard dans le générique parlé du *Roman d'un tricheur* – a toutefois rapidement cédé la place à l'exposition sommaire du récit filmique. Il n'est dès lors pas étonnant que Vinzenz Hediger compte la voix-over (*Sprecherstimme*) parmi les trois traits caractéristiques (avec les mentions verbales et les effets optiques) de la période classique du *trailer* qui se met en place à l'époque de la réalisation de *La Fiancée de Frankenstein*<sup>578</sup>. La forme du premier enchâssement de ce film s'inspire des bandes-annonces qui, comme les prologues, ont connu à la fin de l'interrègne l'éviction de la représentation filmique d'une performance scénique au profit de la voix-narration, c'est-à-dire d'une intégration de tous les fragments en un ensemble cohérent. La voix-over masculine de ce prologue, si elle n'est pas autoritaire comme elle l'est fréquemment dans les documentaires et les bandes-annonces, n'en fait pas moins *autorité* puisqu'il s'agit, dans l'univers diégétique, de la voix du poète Byron, caution littéraire s'il en est. Contrairement aux voix-over totalement désincarnées de la plupart des bandes-annonces, celle-ci est associée à une source diégétique. Situé à une période charnière, *La Fiancée de Frankenstein* prend soin de concilier l'évanescence de la voix-over avec la visualisation des locuteurs.

Alors que la première occurrence *over* de *La Fiancée de Frankenstein* s'opère sur le mode de l'encadrement d'une série d'inserts, la seconde n'implique pas la même présence de l'énonciateur : la voix de Mary Shelley est utilisée pour faire démarrer l'histoire (la phrase « Imaginez-vous face aux décombres du moulin » est indirectement adressée aux spectateurs du film), mais disparaît rapidement au profit de l'audiovisualisation de cette dernière, et ne revient pas dans un épilogue, le film se terminant avec la résolution du récit « enchâssé » dont rien ne vient rappeler le statut (comme cela est également le cas dans d'autres films hollywoodiens ultérieurs, notamment, pour prendre l'exemple d'un personnage féminin, dans *Rebecca* d'Alfred Hitchcock). La voix de la romancière de *La Fiancée de Frankenstein* n'est que très momentanément *over*, se contentant de jouer le rôle de pont sonore entre les deux niveaux de la diégèse. En outre, sa fonction proprement narrative est seulement suggérée par la situation-cadre, car les images prennent le

578. *Ibid.*, p. 118.



relais de la voix avant que n'apparaisse le premier événement, la coréférence verbo-iconique ne portant que sur des «éléments du monde» (le moulin, la cohorte des villageois), non sur des prédicats actionnels. En se limitant à une description de la situation, la voix-over de Shelley est encore majoritairement soumise à cette fonction des prologues et des bandes-annonces qui consiste à mettre l'audiospectateur dans l'atmosphère du film. Elle est à la fois une amorce qui sert à susciter l'intérêt (le pôle de l'attraction) et, virtuellement, le véhicule d'informations narratives.

On peut donc conclure de cette étude que *La Fiancée de Frankenstein*, réalisé à la toute fin de l'interrègne, procède d'un régime vocal intermédiaire entre voix-attraction et voix-narration, partage qui s'opère d'ailleurs au sein même de son prologue entre la femme et l'homme. Le cinéma «classique» des années 1940 et 1950 fera de la voix-over un vecteur essentiel de la narration. La voix-attraction des prologues issus des spectacles de vaudeville ne réapparaîtra en tant que régime dominant que dans des pratiques marginales ou dans un genre comme la comédie musicale.



## VI. VOIX-NARRATION ET ENONCIATION FILMIQUE

### VI.1 LA NOTION DE «VOIX NARRATIVE» DANS LES ETUDES LITTERAIRES

À travers une discussion de la notion de «voix narrative» couramment utilisée par la critique littéraire depuis l'essor de la théorie du roman au XX<sup>e</sup> siècle, j'aimerais montrer en quoi la voix-over est une manifestation emblématique du pôle de l'intégration narrative. A priori, les termes «voix-narration» et «voix narrative» semblent s'appliquer à des phénomènes totalement dissemblables puisque la première se réfère à une occurrence sonore effective, alors que la seconde, issue de l'analyse narratologique des textes littéraires, ne comprend pas le mot «voix» dans son acception concrète. Toutefois, il me semble que les narratologues ne se débarrassent jamais totalement, lorsqu'ils utilisent la notion de «voix», du sens littéral de ce terme qui, comme on le verra, fait retour dans leurs écrits en dépit de sa valeur métaphorique proclamée.

Dans un ouvrage qui posait dans le domaine francophone les bases d'une étude narratologique de la forme romanesque en abordant de façon systématique des questions relatives à la temporalité et au point de vue, Jean Pouillon soulignait combien il était important de faire primer la lecture sur l'audition, tout en donnant une précision où il suggérait une forme particulière d'oralité :

«Si, malgré tout, on veut que l'oreille intervienne, qu'on songe à l'oreille intérieure, celle qui en nous saisit notre langage intime. Je ne veux pas dire par là que le roman doit être nécessairement écrit dans le style du langage intérieur, mais qu'il doit pouvoir s'y intégrer, en suivre le rythme. Lire un roman, c'est entendre quelqu'un nous parler du dedans, et non lire un discours, un exposé.»<sup>579</sup>

579. Jean Pouillon, *Temps et roman*, Paris : Gallimard, 1946, p. 14.

Cette remarque laisse présager des parentés entre la forme romanesque, qui est soumise selon Pouillon à des contraintes formelles relativement faibles (comparativement au discours poétique) et qui, de ce fait, se trouve prédisposée à créer un rapport d'immédiateté avec le lecteur, et le « discours intérieur » d'Eikhenbaum dont on a dit qu'il subissait d'importantes modifications en transitant par la médiation bonimentorielle. En recourant à la notion de « rythme », Pouillon évoque l'adéquation intime qui s'opère entre la production signifiante et son destinataire. « Entendre quelqu'un nous parler du dedans » : l'intuition de Pouillon connaîtra de nombreux avatars chez les théoriciens ultérieurs, quant à eux moins disposés qu'il n'y paraît de prime abord à évacuer la dimension proprement orale de la communication, ainsi qu'en témoigne la fréquente association du narrateur à l'origine d'une voix.

Ce point fait précisément problème lorsqu'on veut appliquer les modèles élaborés pour le roman au cinéma, car l'anthropomorphisation postulée par les théoriciens de la littérature n'est pas appropriée à un tel dispositif « machinique ». Cependant, la représentation cinématographique dominante tente justement d'occulter l'origine technique de sa production par le truchement de la voix humaine. Par conséquent, c'est paradoxalement à travers la conceptualisation du récit scriptural proposée par les partisans de la « voix narrative » que l'on peut s'interroger sur certains principes de l'énonciation filmique.

Même si la théorisation de la voix narrative a permis de prévenir l'assimilation malencontreuse de l'auteur au narrateur, elle demeure sous-tendue par un désir d'individualisation et d'humanisation de la « source » du récit. Bien que cette origine n'y soit pas machinique comme au cinéma, elle se présente avant tout, dans le récit scriptural, sous la forme factuelle du texte. Or l'utilisation du terme « voix », associé à ce qui fait en propre l'identité d'un individu qui s'affirme en tant que « sujet parlant » (en signalant son sexe, son âge, son tempérament, etc.), est symptomatique du caractère anthropocentrique de ce type d'analyse textuelle. On peut dire que la source humaine constitue en quelque sorte le « refoulé » des théories narratologiques qui, principalement dans les années 60 à 80, recoururent à l'analyse de la voix narrative. Considérées sous cet angle, les différentes typologies et hiérarchisations font écran aux motivations premières de cette élaboration théorique. La notion de « voix narrative » est de ce fait « problématique » au sens positif du terme. Il ne s'agit pas de contester ici cet usage

terminologique extrêmement répandu, mais d'analyser ce que « cache » le consensus autour de cette notion et en quoi il nous permet d'interroger deux facteurs couplés depuis les temps immémoriaux des premiers conteurs : la voix et le récit.

L'emploi de la notion de « voix narrative » a été largement popularisé – et cela également parmi les théoriciens du cinéma – par Gérard Genette qui, dans *Figures III*, fait de la « voix » l'une des grandes catégories de l'analyse des récits. Le narratologue motive le choix de ce terme en mentionnant d'une part le sens que lui ont donné les grammairiens, d'autre part, de façon marginale, le titre d'un ouvrage de Marcel Muller qui, à l'instar de sa propre entreprise (mais dans une optique différente, exclusivement « proustologique »), s'est attaché à la *Recherche du temps perdu*<sup>580</sup>. Genette se réfère de façon superficielle à la définition grammaticale, citant à deux reprises la même description sommaire tirée du *Petit Robert* selon laquelle la « voix » consiste en « l'aspect de l'action verbale dans ses rapports avec le sujet »<sup>581</sup>. Ce que le narratologue retient de cette acception des grammairiens – qui regroupent sous cette notion certaines formes verbales (le passif opposé à l'actif), c'est, de façon très générale, la mise en relation d'un sujet et d'un objet. En déplaçant ce principe dans le champ de la narratologie, et donc en substituant au sujet grammatical de l'énoncé une instance d'énonciation (narrative), Genette effectue un transfert qui induit d'importantes différences : en fait, la notion n'entretient plus chez lui qu'une relation « métaphorique » avec la catégorie grammaticale homonyme.

Par contre, la référence à l'étude de Muller sur Proust, c'est-à-dire à une acception de la notion de « voix » qui a traditionnellement cours chez les poéticiens et les théoriciens du roman, permet de mieux appréhender les enjeux de l'approche genettienne. Dans son ouvrage, Muller tente de saisir les niveaux convoqués par le « je » du roman de Proust en ne distinguant pas moins de dix instances différentes. À cette fin, il établit une terminologie spécifique (dont j'épargnerai à mon lecteur tous les détails) où chaque terme se différencie de l'usage commun par l'emploi d'une majuscule. Cette rigueur dans la distinction des instances énonciatives, tout

580. Marcel Muller, *Les Voix narratives dans « La Recherche du temps perdu »*, Genève : Droz, 1965.

581. G. Genette, *Figures III*, *op. cit.*, pp. 75 et 226.

à fait justifiée dans le cas d'une œuvre comme *La Recherche*, amène Muller à formuler le constat suivant: «Héros-je, Héros-nous, Sujet Intermédiaire, Narrateur: telles sont les voix que nous entendons dans le récit.»<sup>582</sup> On pourrait croire que l'expression «entendre des voix» relève, sinon de l'hallucination auditive, du moins d'une formulation strictement métaphorique. Or certains passages suggèrent une référence à la «voix» en tant que phénomène acoustique. Muller écrit par exemple:

«Le Héros ne laisse pas entendre sa voix. Celle qu'entend Picon («un moi... que nous connaissons au timbre de sa voix») est celle, non du Héros, mais du Narrateur, lequel l'emprunte à l'Écrivain. Et aux dires de Paul Morand, la phrase écrite de Proust ressemble étonnamment à sa phrase parlée.»<sup>583</sup>

Une fois reconstituée la chaîne des instances, on aboutit, en dépit des distinctions méthodologiques, à ce fait élémentaire: il y a quelque chose dans le texte qui relève de la «voix» de Proust, au sens littéral d'un «son créé dans le larynx par les vibrations des ondes vocales au passage de l'air provenant des poumons»<sup>584</sup>. Et les «poumons» de l'œuvre sont bien sûr ceux de l'Auteur, auquel Muller se réfère dans une hypothétique situation d'oralité. Le postulat implicite de Muller est le suivant: Marcel Proust écrit comme il parle, c'est pourquoi le terme «voix» s'avère pertinent pour commenter son texte.

Autant l'expression «point de vue» est entachée de son ancrage dans le regard (c'est pourquoi Genette lui préférera la notion de «focalisation», qui ne pose pas moins de problèmes lorsqu'on l'associe dans les études cinématographiques à la «focale» de la caméra), autant celui de «voix narrative» semble constamment renvoyer à une situation d'oralité. Quinze ans après l'ouvrage de Muller, Jean Peytard utilise le cadre théorique de la «voix narrative» pour examiner la nouvelle *Vanina Vanini* de Stendhal, et commence par affirmer: «Si la lecture choisit d'écouter-voir (la) (les) voix narrative(s) qui module(nt) le récit, c'est qu'elle dépend du simple postulat que Stendhal écrit.»<sup>585</sup> La dimension sonore

582. M. Muller, *art. cit.*, p. 19.

583. *Ibid.*, p. 16.

584. Premier sens du terme donné in Bernard Pottier (dir.), *Le Langage*, Paris: Centre d'étude et de promotion de la lecture, 1973, p. 538.

585. Jean Peytard, *Voix et traces narratives chez Stendhal*, Paris: Les Éditeurs Français réunis, 1980, p. 10.

demeure sous-jacente, et l'hésitation entre le singulier et le pluriel trahit l'existence de l'unité du sujet parlant par-delà ses avatars fictionnels (le narrateur comme personnage) et fictifs (le narrateur comme instance d'énonciation).

Ce qui est (à la rigueur) valable pour la littérature ne l'est par contre aucunement pour le cinéma, où le film résulte d'un travail collectif. Cependant, comme les films parlants proposent concrètement des voix (et notamment celles de narrateurs), on peut s'interroger sur les raisons de cette connotation d'oralité dont les analyses littéraires ne se défont pas, en dépit de l'affirmation réitérée de la part de leurs auteurs que le terme «voix» doit être compris comme une notion abstraite. En quoi ce que nous avons dit de la voix vive innerve-t-il certaines conceptions narratologiques? Si l'on prend par exemple l'une des «anciennes» tentatives de systématisation, celle de William C. Booth, on constate à la seule lecture de la table des matières l'importance octroyée au champ lexical de la «voix» pour aborder des questions narratologiques impliquant également l'«auteur». En effet, l'une des trois grandes parties s'intitule «La voix de l'auteur dans la fiction» (fût-il «implicite», cet «auteur» est bien comparé à un sujet parlant), et l'ouvrage comprend de surcroît le chapitre «Les usages du *silence* auctorial» et la section «Les nombreuses voix de l'auteur»<sup>586</sup>. Nonobstant la distinction productive que Booth opère entre «auteur implicite» et narrateur, il soumet sa typologie à un dessein plus général qui consiste, ainsi que le résume Françoise Van Rossum-Guyon, à s'élever «contre le mythe de la disparition de l'auteur»<sup>587</sup>.

Si la notion de «voix» comprise au sens d'une «production sonore» continue de sourdre dans ces écrits, c'est selon moi principalement pour deux raisons, dont l'examen m'apparaît fort éclairant si l'on veut saisir les limites de la pertinence de ces théories lorsqu'on les applique au cinéma. D'une part, la théorie de l'énonciation narrative s'appuie sur un modèle communicationnel – l'apport essentiel de cette approche réside d'ailleurs dans le fait qu'elle a permis d'élargir l'objet d'étude au-delà de «l'immanence du texte» des structuralistes – qui ne peut se passer de la situation courante dont il est issu, celle de la conversation orale entre *individus*, et donc

586. William C. Booth, *The Rhetoric of Fiction*, Chicago: University of Chicago Press, 1961.

587. Françoise Van Rossum-Guyon, «Point de vue ou perspective narrative», *Poétique*, N° 4, 1970, p. 482.

d'une origine *personnalisée* du verbal. Pour S.-Y. Kuroda, dont l'article écrit au milieu des années 70 porte sur la décennie précédente où fleurirent de nombreuses propositions de cet ordre, «il n'est pas étonnant que de nombreuses théories modernes de la narration soient, explicitement ou implicitement, sous l'influence de la théorie communicationnelle de la performance linguistique»<sup>588</sup>. Ce rejet dans l'implicite (quand il ne s'agit pas d'un impensé théorique) est d'autant plus fort dans le champ des études sur le cinéma que le médium exige de prendre ses distances face au modèle linguistique. Dans les analyses du texte littéraire, le phénomène de «présence-absence» induit par la voix séparée du corps d'un locuteur – l'écrit se désolidarisant de l'écrivain du fait même de l'écriture – permet de rendre compte des «traces» de la subjectivité d'une instance narrative qui est à la fois «perçue» par le lecteur et évincée de la transmission effective du récit. Or nul mieux que le statut *over* ne permet à la voix filmique, pourtant phénomène purement acoustique, de tendre à l'abstraction de la voix narrative. En niant la visualisation de sa source dans un médium défini par l'image, la voix-over constitue un moyen privilégié pour incarner, grâce à la désincarnation même qui la définit, ce narrateur insaisissable que Todorov qualifie «d'image fugitive [qui] ne se laisse pas approcher»<sup>589</sup>. Peytard évoque à propos du récit littéraire «ce qui marque le récit comme un acte de langage partagé, et témoigne que la voix narrative est la trace du désir d'être écouté, d'être lu», la lecture n'apparaissant qu'en seconde position alors qu'il ne devrait être question que d'elle<sup>590</sup>. Le Français subit probablement l'influence du contexte d'origine de son approche, celui de la théorie d'Émile Benveniste. En effet, lorsque ce dernier aborde la question de la «subjectivité dans le langage» ou définit la catégorie du «discours», il s'appuie sur une situation dialogique orale. Il précise par exemple que «la distinction discours *vs.* histoire ne recoupe pas celle que l'on peut établir entre l'oral et l'écrit, mais il est évident que la situation orale fait office de modèle pour appréhender la catégorie du discours». Ainsi mentionne-t-il comme objet de recherche autant les situations conversationnelles que les

588. Sige-Yuki Kuroda, «Réflexions sur les fondements de la théorie de la narration», in Julia Kristeva, Jean-Claude Milner, Nicolas Ruwet (dir.), *Langue, discours, société*, Paris: Seuil, 1975, p. 261.

589. Tzvetan Todorov, *Littérature et signification*, Paris: Larousse, 1967, p. 87.

590. J. Peytard, *op. cit.*, p. 88.



«écrits qui reproduisent des discours oraux ou qui en empruntent le tour et les fins»<sup>591</sup>. Certes, Benveniste conçoit l'«énonciation historique» comme un masquage du contexte communicationnel qui résulte de l'effacement de toute marque énonciative; néanmoins, comme il définit l'«histoire» par la négative, tout ce qu'il en dit a été reporté par ses continuateurs sur la notion-clé de «discours». Or le texte littéraire, à l'instar du texte filmique, n'offre aucune possibilité d'interactivité. L'impossible transformation du «tu» en «je», voire l'absence d'équivalents de tels embrayeurs au cinéma ne pose pas particulièrement problème<sup>592</sup>, d'une part parce que certaines situations d'oralité peuvent être mises en scène à l'intérieur même du film, d'autre part parce que les catégories du discours et de l'histoire valent surtout en tant que principes généraux d'interaction entre le texte et son destinataire. On constate néanmoins que les références à l'oralité (caractérisée par une relation directe entre l'énonciateur et l'énonciataire) subsistent même lorsqu'il est question de *média*. À l'instar de la voix enregistrée, le texte scriptural se présente, du moins depuis la fin du XII<sup>e</sup> siècle<sup>593</sup>, comme une entité qui, dans l'esprit du lecteur, est à la fois indépendante de son support matériel et séparée de tout acte de production. C'est pourquoi le lecteur tend à reconstruire un rapport d'immédiateté avec la source en se figurant qu'un espace-temps de la production est inscrit en creux dans le texte. Paul Ricoeur, dont l'entreprise théorique tente de corrélérer le temps de la mise en intrigue avec celui de l'expérience du lecteur, est très attentif à cette problématique:

«Qu'en est-il de la voix narrative? Cette catégorie littéraire ne saurait être éliminée par celle de point de vue, dans la mesure où elle est inséparable de celle, inexpugnable, de narrateur, en tant que projection fictive de l'auteur dans le texte lui-même. [...]. Le narrateur – locuteur de la voix narrative – ne peut être affranchi au même degré de toute métaphore personnalisante, dans la mesure où il est auteur fictif du discours.»<sup>594</sup>

591. Émile Benveniste, «Les relations de temps dans le verbe français», in *Problèmes de linguistique générale*, Vol. 1, Paris: Gallimard, 1966, p. 242.

592. Voir Ch. Metz, *L'Énonciation impersonnelle*, op. cit., pp. 13-14.

593. Voir Ivan Illich, *Du lisible au visible: la naissance du texte*, Paris: Éditions du Cerf, 1991.

594. Paul Ricoeur, *Temps et récit*, tome 2, Paris: Seuil, 1984, pp. 181-182.

La nécessité d'une telle personnalisation paraît néanmoins discutable au cinéma – comme elle l'est, dans une moindre mesure, dans l'analyse littéraire (les notions de « personne » et de « subjectivité » ayant chez Benveniste un sens purement linguistique) – puisque l'instance responsable de l'organisation filmique ne peut être comparée, même par le plus convaincu des partisans de la « politique des auteurs », à ce que Ricoeur qualifie pour le texte littéraire de « projection fictive de l'auteur ».

Une ressemblance *illusoire* se dessine pourtant au travers d'un procédé spécifique: la voix-over. Ce n'est en effet pas un hasard si Henderson, dans un article où il applique au cinéma les notions proposées dans l'ouvrage de Genette (paru trois ans auparavant dans sa version anglaise), se consacre entièrement à la narration en voix-over dès lors qu'il s'agit d'illustrer la catégorie de la « voix »<sup>595</sup>. Il y a là une congruence qu'il me paraît important d'examiner à partir d'un théoricien, Albert Laffay, qui, bien que travaillant de manière empirique, a avancé dans le champ du cinéma, quelque vingt-cinq ans avant Gérard Genette, certaines hypothèses sur la source de l'énonciation filmique qui ont été à mon sens trop peu discutées.

## VI.2 LA « REVELATION » D'ALBERT LAFFAY

En 1947, Albert Laffay, que l'on peut considérer comme le précurseur des réflexions sur le concept de « narrativité » au cinéma qui seront développées plus tard dans la foulée poststructuraliste<sup>596</sup> et, surtout, sur ce qui deviendra « l'énonciation filmique », faisait le constat suivant: « Il y a toujours derrière un film une sorte de récitant en puissance. »<sup>597</sup> Bien que la seule chose dont les successeurs de Laffay paraissent se souvenir soit la notion de « Grand Imagier », qui renvoie a priori exclusivement à la dimension visuelle, Laffay

595. Brian Henderson, « Tense, Mood, and Voice in Film (Notes after Genette) », *Film Quarterly*, Vol. 36, N° 4, 1983.

596. Voir A. Gaudreault, « Les aventures d'un concept: la narrativité », in Michel Marie et Marc Vernet (dir.), *Christian Metz et la théorie du cinéma*, Paris: Méridiens Klincksieck, 1990, pp. 121-122.

597. Albert Laffay, « Le récit, le monde et le cinéma », *Les Temps modernes*, N° 21, mai-juin 1947, pp. 1597-1598. Cet article est repris intégralement, à quelques changements près, dans le troisième chapitre de *Logique du cinéma*, notamment dans la section significativement intitulée « Le montreur d'images » (*Logique du cinéma, op. cit.*, pp. 72-83). Je me réfère néanmoins à la version des *Temps modernes* parue dix-sept ans plus tôt parce que certaines remarques sur l'époque contemporaine de l'écriture de la première version, intéressantes du point de vue de l'étude de la voix-over, ont subi des changements lors de l'édition du livre.

accorde en fait, si l'on y regarde de plus près, une grande importance à l'introduction d'un «récitant» dans le film. De toute évidence, c'est même la voix-over qui lui a fait prendre conscience de la présence d'une instance qui serait à l'origine du discours des images, c'est-à-dire de la possibilité d'une véritable narration cinématographique dont il doutait, dans un premier temps, de l'existence. La mise en place d'un récit exige selon lui de se dégager de la concrétude et de l'absence d'ancrage temporel chevillées à la nature de l'image ciné-photographique.

Que la voix-over ait véritablement fait office pour Laffay d'instrument herméneutique ou qu'il la convoque à des fins argumentatives, il est évident qu'elle joue chez lui un rôle décisif. Il faut en effet noter qu'il développe sa théorie dans le contexte de l'immédiat après-guerre, lorsque les productions américaines de type «films noirs», souvent pourvues de voix-over, font leur apparition avec quelques années de retard sur le marché français<sup>598</sup>. Jean-Pierre Chartier, dont le texte sur les films «à la première personne» paraît la même année, se trouve semblablement stimulé par la vision de ces films en voix-over, procédé qu'il dit observer dans «la moitié des grands films américains» contemporains et qui permet selon lui des phénomènes d'intériorisation encore inédits au cinéma. Il souligne notamment, au cours d'une comparaison entre film et roman, l'importance qu'acquiert l'oralité lorsqu'il s'agit d'exprimer des pensées: «Sans doute les mots au cinéma sont parlés au lieu d'être écrits. Mais quand nous essayons d'imaginer notre monologue intérieur, il nous semble plus proche du langage parlé que du langage écrit.»<sup>599</sup> Chartier défend donc une conception identique à celle que nous soutenons ici: la voix-over est généralement aux antipodes de la distanciation, puisqu'elle fait au contraire fusionner le «je» du personnage-narrateur avec le spectateur, renforçant ainsi l'«illusion de réalité». Le fait que cette réflexion qui se verra prolongée chez les filmologues ait été suscitée par la voix-over témoigne du rôle-clé que ce procédé a joué dans les théories du cinéma.

598. L'expression «noirs» pour parler de cette production hollywoodienne est introduite par Nino Frank dans un article paru un peu moins d'une année avant celui de Laffay. Frank fait de l'usage de la voix-over une spécificité de ce type de films, et considère ce procédé comme une «évolution interne» du genre policier (Nino Frank, *art. cit.*, p. 14).

599. Jean-Pierre Chartier, «Les «films à la première personne» et l'illusion de réalité au cinéma», *La Revue du Cinéma*, N° 4, 1947, p. 32.

Laffay souligne également l'importance des films hollywoodiens en voix-over, mais fait quant à lui référence à une forme de « discours intérieur » qui est plus proche de l'acception d'Eikhenbaum : « Or il vient juste de nous arriver d'Amérique tout un lot de films où ce récit virtuel est au moins en partie tiré au grand jour. [...] Je pense naturellement aux bandes où un commentaire parlé accompagne, soutient ou fait naître les images (*Qu'elle était verte ma vallée*, *Citizen Kane*, *Assurance sur la mort*, etc.). »<sup>600</sup> Ainsi la voix-over actualiserait-elle un « discours » sous-jacent qui, sans elle, ne serait pas perçu. Laffay évoque par la suite le rôle de « révélation » qu'a joué pour sa conceptualisation théorique l'utilisation de la voix-over par plusieurs films contemporains :

« Peut-être cette technique n'est-elle qu'une mode. Les films avec récitant nous paraîtront-ils désuets et un tantinet ridicules d'ici quelques années ? Je n'en sais rien. Il se peut bien aussi qu'ils indiquent une direction où le cinéma va s'engager résolument. Ce style, quoi qu'il en soit, aura été le révélateur chimique d'une présence virtuelle cachée derrière *tous les films*, celle d'une sorte de maître de cérémonie, le grand imagier qui donne pour nous aux vues photographiques le sens, le rythme et la durée. »<sup>601</sup>

L'auteur constate une prédisposition de la voix-over à agir comme un « révélateur chimique » de l'énonciation filmique. Bien que le modèle énonciatif que je proposerai ici soit fondamentalement différent des hypothèses émises par Laffay, il me paraît nécessaire de considérer ses remarques comme l'expression d'un fait de réception qui, en lui-même, est « révélateur » d'un certain positionnement du spectateur par les films à voix-over. En effet, la terminologie de Laffay est symptomatique d'une opposition entre la « surface » (ce que la représentation donne à voir et à entendre) et la « profondeur » (les processus cachés de production du sens) que l'on retrouvera, à un autre niveau (celui de la « méga-énonciation »

600. A. Laffay, « Le récit, le monde et le cinéma », *art. cit.*, p. 1592. La phrase est modifiée comme suit dans *Logique du cinéma* : « Or il nous est arrivé pour la première fois d'Amérique aux alentours de 1947 tout un lot de films où ce récit virtuel était au moins en partie tiré au grand jour » (*Logique du cinéma, op. cit.*, p. 77).

601. A. Laffay, « Le récit, le monde et le cinéma », *art. cit.*, p. 1596. Dans l'ouvrage, Laffay modifie sensiblement la formulation, ce qui donne une sorte de réponse à son interrogation de la fin des années 1940 : « Cette technique fut un peu une mode. Les films avec récitant sont peut-être moins répandus aujourd'hui » (*Logique du cinéma, op. cit.*, p. 81).

au sens de Gaudreault), dans les écrits de théoriciens ultérieurs, notamment chez Christian Metz. Ce dernier déclare par exemple à propos de la pratique cinématographique dominante qu'elle repose sur un principe consistant à « effacer les marques d'énonciation et à se déguiser en histoire »<sup>602</sup>. David Bordwell a montré que cette conception d'un discours « sous » l'histoire – que Gérard Genette me paraît avoir contribué à populariser en considérant celle-ci comme une sous-catégorie de celui-là<sup>603</sup> – ne respecte pas les définitions de Benveniste pour lequel ces deux concepts entretiennent une relation d'exclusion mutuelle<sup>604</sup>. La reformulation metzienne me paraît néanmoins pertinente si on ne la prend pas au pied de la lettre, et qu'on l'applique à l'occultation, particulièrement évidente au cinéma, de la partie machinique du dispositif. Pour Laffay, la voix-over n'exhibe pas la technologie sonore qui la produit (il ne discute nullement le primat du synchronisme audiovisuel), mais la structure énonciative inhérente au récit filmique. Chez lui, toute manifestation verbale est instinctivement rattachée à un locuteur. Cette conception, bien que contestable, s'explique aisément. Les spectateurs savent qu'une « machine » peut produire une image sans qu'aucun facteur humain (ou presque) n'intervienne dans sa constitution : les simples orifices ou miroirs de la *camera oscura* l'ont prouvé. Le monde objectal lui-même s'offre à l'observateur sous une apparence visuelle, et les sons de la nature ne s'articulent pas en un langage, en dépit de toutes les utopies des poètes. Par contre, il n'y a pas de paroles sans un être vivant pour les proférer. André Gaudreault a incidemment souligné les implications énonciatives de ce rapport entre l'humain et la machine qui fonde l'existence du cinéma parlant :

« Lors de la communication d'un récit scriptural (en fait, de tout récit verbal), il y a toujours « quelqu'un » qui parle : le narrateur *fondamental*. Qu'il prenne ou non l'oripeau d'un personnage-narrateur (ou même d'un personnage tout court), c'est lui le « moulin » à paroles du récit. Mais que cette image non anthropomorphe, proprement « machinique », du moulin ne porte pas à confusion : le langage apparaît toujours comme le produit d'un sujet humain, son utilisation est très difficilement transparente. »<sup>605</sup>

602. Ch. Metz, *Le Signifiant imaginaire*, op. cit., p. 113.

603. Voir G. Genette, « Frontières du récit », *Communications*, N° 8, 1966, repris in *Figures II*, Paris : Seuil, 1969, p. 66.

604. D. Bordwell, *Narration...*, op. cit., p. 23.

605. A. Gaudreault, *Du littéraire...*, op. cit., p. 95.

À l'époque où le bonimenteur commentait les images dans la salle, on a vu que le bruit du « moulinet » de l'appareil de projection activé par l'opérateur (ou du « moulin » des appareils à entraînement électrique) se faisait entendre simultanément à la manifestation humaine de celui qui commentait les images. Avec les différentes techniques de synchronisation de l'époque « muette », l'émission des voix était également produite par le mouvement mécanique de la rotation d'un cylindre ou d'un disque. Aujourd'hui, elle résulte généralement du défilement même de la pellicule. C'est pourquoi, même lorsqu'il se manifeste verbalement en déléguant fictivement sa fonction à une instance diégétique, le « narrateur filmique » doit son existence à la technologie phonographique, c'est-à-dire au pôle machinique du dispositif cinéma. À la fin du présent chapitre, je tenterai de prendre en compte les implications théoriques de cette donnée fondamentale. Dans un premier temps, je propose de développer certaines intuitions de Laffay en considérant les propositions de ses successeurs. Cette approche me permettra de revenir sur ce qui oppose la voix-narration à la voix-attraction du point de vue de la « présentification », cette fois en abordant la voix-narration dans ses qualités propres.

### VI.3 LA FICTION ASSOCIEE AU PASSE DE LA VOIX (NARRATIVE)

Les théoriciens du roman et les romanciers ont traditionnellement utilisé la notion de « narrateur » pour qualifier cette « présence » diffuse que la manifestation verbale « appelle » au niveau textuel. Laffay se montre particulièrement sensible à cette présence humaine et l'inscrit notamment dans une filiation pré-cinématographique en affirmant à propos de *Qu'elle était verte ma vallée* que « le montreur de lanterne magique y va de son boniment »<sup>606</sup>. Or cette instance implique non seulement un espace virtuel (celui, abstrait, qu'occupe le site désincarné dont émane la voix-over), mais aussi et surtout une certaine temporalité : la présence se perçoit au *présent*. Certes, on l'a dit, la voix-narration ne possède pas la force de présentification de la voix-attraction. Néanmoins, par contraste avec les seules images qu'elle rejette dans le passé, l'énonciation verbale induit une forme de contemporanéité, celle du

606. A. Laffay, « Le récit, le monde et le cinéma », *art. cit.*, p. 1592.

«maintenant» de son émission vocale. Laffay analyse avec beaucoup de clairvoyance cette incidence de la voix-over sur la perception des images :

«Ce qu'il subsiste toujours dans les clichés de nécessité mécanique tend à faire d'un film une succession de simples «présences», contemporaines du monde et de nous autres ses prisonniers, au lieu que *la fiction récitative* lui donne des ailes, *mais le rejette aussi dans l'immuable du passé.*»<sup>607</sup>

Cette explication me paraît tout à fait intéressante, d'abord parce qu'elle oppose la mécanique à la voix-over, ensuite parce qu'elle associe passé et fiction. Sur ce second point, Laffay peut être rattaché à une tradition de linguistes et de théoriciens de la littérature qui ont discuté les raisons des accointances entre la voix narrative et le passé, voire la parenté entre passé et fiction. Le consensus recueilli par les théories issues de champs différents autour de la valeur passée des récits m'incite non pas à leur accorder un crédit sans réserve (elles procèdent souvent d'une confusion regrettable entre récit et fiction), mais à les considérer comme une donnée historiographique d'importance. Pour Laffay, le spectateur prend son «envol» lorsque la voix-over le transporte dans le passé, éloigné de lui comme le sont les univers fictionnels. Les liens entre la «fiction récitative» et le passé s'expliquent à mon sens par la nature verbale du texte *over* qui dispose de temps verbaux, alors que l'image est invariablement liée à l'ici-maintenant de sa projection. La voix désincarnée se trouve doublement associée au passé, tant sur un plan phylogénétique (la tradition de l'oralité liée aux récits des origines) qu'ontogénétique (la salle de cinéma comme espace intra-utérin, la voix-over comme extériorisation du «discours égocentrique» au sens de Piaget et Vygotski). Pourtant, si la voix peut s'imposer en tant que véhicule de la transmission d'un récit qu'elle rejette dans le passé, c'est qu'elle bénéficie d'un ancrage dans le présent. On peut alors transposer la valeur spatiale de la «présence-absence» qui concerne doublement la voix-over (en tant que voix phonographiée et acousmatique) dans le domaine temporel : la voix est à la fois «au présent» en raison de sa matérialité vocale et «au passé» parce qu'elle n'est que la trace d'une émission sonore désormais disparue.

607. *Ibid.*, pp. 1597-1598 (je souligne).

Sur la base de cette conversion de la distance spatiale en distance temporelle s'échafaude à mon sens le pacte fictionnel. En vertu du paradoxe de la présence-absence qui caractérise toute voix enregistrée et plus particulièrement la voix-over, celle-ci se prête particulièrement bien à l'introduction de *flash-backs*. Or, souvent inaugurale, elle invite le spectateur dans la fiction du film en se faisant le pivot d'un basculement dans le passé. Le rapprochement entre passé et fiction, s'il a une réalité du point de vue de l'historiographie des théories narratologiques, doit toutefois être nuancé : la distinction entre fiction et non-fiction ne peut s'établir que sur des bases pragmatiques, grâce à l'intervention du paratexte, et nullement à partir d'indices purement textuels. Le primat de la dimension pragmatique légitime par contre l'examen de ces théories comme des *faits de réception* : s'il apparaît qu'une communauté de spectateurs a tendance à associer passé et fiction, ce lien devient opératoire, c'est-à-dire qu'il peut être utilisé par les films comme un indice de fictionalité.

Les films documentaires, susceptibles de présenter un large éventail de voix, comprennent fréquemment une voix-over et recourent également aux temps passés<sup>608</sup>. Ce fait interdit que la valeur prétéritive constitue en elle-même un facteur discriminant. Cependant, dans de tels films, la voix ne conte pas les événements passés en oblitérant leur possible impact sur le monde d'expérience du spectateur, mais souligne au contraire la distance temporelle qui les sépare du « maintenant » de la réalisation (ou de la projection) du film. Ainsi, la lecture du spectateur est généralement orientée afin qu'il considère l'image comme un « ça a été », le statut de réalité reposant justement sur un ancrage dans le passé. Dès lors, les temps passés du texte *over* des films documentaires conservent une valeur temporelle effective et peuvent ainsi instituer une véritable distance, mais aussi conforter la croyance en l'existence réelle des événements montrés. Dans un film comme *Nuit et brouillard* (Alain Resnais, 1955) qui se présente en tant que témoignage et exige par conséquent que le référent de la représentation soit considéré comme réel, l'opposition entre les phrases au présent (par exemple « Au moment où je vous parle ») accompagnant les images en couleurs et celles conjuguées au passé sur les images d'archives en

608. Cf. F. Niney, *op. cit.*, pp. 236-237.



noir/blanc participe d'une démarche globale d'authentification du montré<sup>609</sup>.

Par contre, la césure franche entre le monde «réel» et l'univers diégétique fait partie du pacte fictionnel, en rapport inverse avec le degré de réalisme postulé: l'éviction du repère qu'est le présent agit sur la suspension de l'«incrédulité» du spectateur. C'est pourquoi la question «énonciative» de l'association entre passé et fiction proposée par Laffay permet d'aborder l'une des fonctions fondamentales de la voix-over, celle qui consiste à «créer un monde».

Dans le champ des études cinématographiques, on retrouve notamment l'idée de Laffay au début des années 80 chez Henderson, pour qui la fonction principale de la voix-over réside, du moins dans le «cinéma classique», dans le fait de ménager une transition entre le présent et le passé en exploitant l'univocité offerte par les temps verbaux<sup>610</sup>. Il s'agit d'une fonction qu'un grand nombre de théoriciens du roman ont également rapportée à la voix narrative. Depuis Käte Friedemann qui, dans les années 10, compare drame et récit épique, on trouve par exemple la discussion du passé d'événements racontés par une instance de médiation inscrite dans un présent chez Stanzel ou Hamburger (pour s'en tenir aux auteurs germaniques). Il s'agit même d'une des thèses centrales de *Temps et récit* de Ricœur, qui la synthétise dans le troisième tome en rappelant que «les événements racontés dans un récit de fiction sont des faits passés pour la *voix narrative* que nous pouvons tenir ici pour identique à l'auteur impliqué, c'est-à-dire à un déguisement fictif de l'auteur réel»<sup>611</sup>. Or cette opposition entre présent et passé, qui n'a de sens qu'au sein du couple temps diégétique/temps de l'énonciation (temps du texte/temps de l'acte selon la terminologie de Weinrich), trouve dans le cinéma parlant une correspondance: il s'agit de la distinction entre la perception par le spectateur du caractère construit de l'image et l'immédiateté présumée des sons, dont la nature d'enregistrement phonographique est bien plus aisément occultée que, par exemple, la bidimensionnalité de la représentation écranique.

La question de la temporalité de l'acte narratif entre étonnamment en résonance avec ce que Sacha Guitry disait du cinéma par opposition au théâtre (rejoignant une comparaison que l'on trouve

609. Voir A. Boillat, *La Fiction au cinéma*, op. cit., pp. 91-95.

610. B. Henderson, art. cit., p. 16.

611. P. Ricœur, *Temps et récit*, tome 3, Paris: Seuil, 1985, p. 276.

entre autres chez Friedemann, Hamburger et Gaudreault): au cinéma, «l'acteur *a joué*» (voir chapitre III). C'est d'ailleurs précisément dans le contexte de l'énonciation au théâtre qu'André Gaudreault évoque le cas d'un «personnage en voix *off*» pour mettre en évidence l'instance qui est à la source du récit scénique (le monstreur): «C'est peut-être dans de telles occasions d'ailleurs que le *monstreur* arrive à donner «de lui-même», de la manière la plus stricte, une image de *narrateur*, par une imitation *orale* (donc verbale) de l'ancrage *scriptural* (donc verbal) de la voix narrative.»<sup>612</sup> La «voix-narrative» est donc intimement associée à *une* écriture. Ce cadre théorique permet de comprendre pourquoi Guitry, en dérivant l'esthétique du *Roman d'un tricheur* du constat d'un temps filmique fondamentalement passé, aboutit à un usage radical de la voix-over. En supprimant toutes les autres voix, il dénie au monde audiovisuelisé un possible ancrage dans le présent. Par contre, cette voix unique s'origine fortement dans une situation-cadre «au présent» qui réapparaît ponctuellement.

La réflexion sur le passé de la voix narrative ne s'est pas limitée, comme on l'a dit, à la question de la temporalité, mais a été élargie à la fiction. La thèse qui est à cet égard la plus originale est celle que Käte Hamburger formule à propos du temps verbal de l'imparfait (c'est-à-dire du prétérit allemand, qui comprend également le passé simple français). Selon elle, dans la fiction (c'est-à-dire lorsque l'énonciateur réel devient une «Je-origine fictive»), le passé perd sa valeur prétéritive pour devenir l'un des indicateurs du statut fictionnel des événements contés<sup>613</sup>. En dépit de confusions évidentes (entre fiction et littérature, et de ce fait entre cinéma fictionnel et littérature), la thèse de Hamburger m'a toujours paru avoir un fond de vérité (comme on dit qu'une voix nous parvient du «fond des âges»), sans que je puisse véritablement expliquer pourquoi: c'est bien, me semble-t-il, que les processus décrits s'enracinent dans la réalité subjective de la réception des récits. Cette atemporalité du régime fictionnel est manifeste dans l'expression «Il était une fois», à propos de laquelle Weinrich avait déjà noté qu'elle nous introduit à «un univers régi par un Temps propre, n'ayant qu'une ressemblance fort lointaine avec celui qu'indiquent les horloges»<sup>614</sup>.

612. A. Gaudreault, *Du littéraire...*, *op. cit.*, p. 93.

613. Käte Hamburger, *Logique des genres littéraires*, Paris: Seuil, 1986 [1977], p. 77; voir également A. Boillat, *La Fiction au cinéma*, *op. cit.*, pp. 69-71.

614. Harald Weinrich, *Le Temps*, Paris: Seuil, 1964, p. 46.

L'une des manifestations de cette thèse sur le passé me paraît être l'utilisation de l'imparfait que font les enfants dans leurs jeux «fictionnels». Dans une analyse grammaticale de l'utilisation de l'imparfait français à l'intérieur des phrases hypothétiques, Jean-Michel Adam propose de rapprocher le mode conditionnel de la construction d'un monde. Alors que la fiction repose sur un «comme si», c'est justement à l'imparfait que la langue française recourt dans de tels énoncés. Parmi les usages «marginaux» de l'imparfait qui tendent à révéler sa fonction d'«opérateur de fictionalité», Adam mentionne l'«imparfait préludique» utilisé par les enfants en situation de jeu<sup>615</sup>. Sur ce point, il est intéressant de noter l'intuition suivante d'Albert Laffay concernant le «passé de la fiction» :

«Autrement dit, on passe directement du passé au futur ou, pour parler plus exactement, il n'y a au fond que du passé fictivement orienté vers l'avenir. Le futur d'un récit est un futur imaginaire. Ce qui me fait penser à ces jeux d'imitation des enfants où tout demeure au conditionnel: «Tu serais la maman, j'irais te faire une visite.»<sup>616</sup>

On peut en effet prolonger le constat de la suspension d'une inscription temporelle effective – que Laffay rapporte à tous récits, alors qu'il ne peut s'agir que de récits fictionnels – à la corrélation étudiée par Jean-Michel Adam entre conditionnel et imparfait. Cette réflexion sur la temporalité des récits trouve un écho dans les travaux plus récents du philosophe Paul Ricœur. À partir de l'hypothèse formulée par Käte Hamburger de l'absence totale de valeur temporelle des temps passés dans la fiction et celle, moins radicale, d'un Weinrich qui parle de la suspension de l'engagement du lecteur dans son environnement réel, Paul Ricœur envisage une solution intermédiaire qui permet de concilier valeur passée et fictionalité. Il développe l'idée d'un «quasi-passé» et postule que l'entrée en lecture – que l'on peut appliquer à tous types de mondes, y compris ceux produits par une représentation audiovisuelle – consiste à «inclure dans le pacte entre le lecteur et l'auteur la croyance que les événements

615. Jean-Michel Adam, «Si hypothétique et l'imparfait», in *Langue et littérature*, Paris: Hachette, 1992, pp. 75-78.

616. A. Laffay, *Logique du cinéma*, op. cit., p. 20.

rapportés par la voix narrative appartiennent au passé de cette voix»<sup>617</sup>. La nature contractuelle de cette croyance nous rapproche du fonctionnement pragmatique de la fiction; Ricœur fait d'ailleurs le pont entre passé et fiction en déclarant que le «quasi-passé» conduit à suspendre «la croyance au passé comme ayant été, pour la transposer au niveau de la fiction»<sup>618</sup>.

Le phénomène de collusion entre passé et fiction décrit par Ricœur me paraît essentiel pour saisir les conséquences de l'assimilation de la voix-over des récits rétrospectifs à la voix narrative d'une part, à une composante fictionnelle d'autre part. Grâce aux temps verbaux, la voix-over permet de définir précisément une certaine temporalité, l'image se caractérisant, comme l'ont montré les sémiologues, par un perpétuel présent<sup>619</sup>. Ainsi la représentation visuelle se voit-elle happée par l'ancrage temporel de la voix-over qui lui assigne, généralement, une place dans le passé. Or les implications fictionalisantes de ce rejet des événements (audio)visualisés dans le passé donnent à penser que la voix génère la fiction. C'est pourquoi la voix-over, dont l'un des «ancêtres» est à cet égard le prologue filmique, se manifeste très souvent dès le tout début des films, parfois même avant la première image<sup>620</sup>. Christine Brinckmann a très justement qualifié cette voix-over d'introduction de «signal fictionnel»<sup>621</sup>.

L'image et le texte de la séquence inaugurale de *Qu'elle était verte ma vallée* (John Ford, 1941) illustrent de façon presque archétypale la fonction de création d'un monde qu'endosse le passage du présent au passé lorsqu'il est médiatisé par une voix-narration. D'ailleurs, le «prétérit épique» de Hamburger apparaît dès le titre du roman de Llewellyn repris pour le film. Un texte *over* au présent se réfère tout d'abord à la situation que nous voyons à l'image,

617. P. Ricœur, *Temps et récit*, tome 3, *op. cit.*, p. 276.

618. P. Ricœur, *Temps et récit*, tome 2, *op. cit.*, p. 139.

619. Voir Roger Odin, *Cinéma...*, *op. cit.*, pp. 72-73.

620. Dans le récent *Doom* (Andrej Bartowiak, 2005) adapté d'un jeu vidéo, la voix-over féminine inaugurale qui présente quelques éléments du monde futuriste du film intervient sur fond noir avant même l'apparition du logo de la Universal. Au sein de la production classique hollywoodienne, la position liminaire des voix-over est très fréquente. Le début constitue parfois l'unique lieu d'une manifestation *over* de tout le film. On peut penser à des réalisations aussi différentes en termes de genre que *Rebecca* (Alfred Hitchcock, 1940), *Jeux dangereux* (*To Be or not To Be*, Ernst Lubitsch, 1942), *La Maison de la 92<sup>e</sup> Rue* (*House on the 92th Street*, Henry Hathaway, 1945), *Vaudou* (*I Walked with a Zombie*, Jacques Tourneur, 1947), *La Flèche brisée* (*Broken Arrow*, Delmer Daves, 1950), *Ils étaient des hommes* (*The Men*, Fred Zinneman, 1950), *Tous en scène* (*The Band Wagon*, Vincente Minnelli, 1953) ou *Cléopâtre* (*Cleopatra*, Joseph L. Mankiewicz, 1963).

621. Christine N. Brinckmann, «Der Voice-Over im Film Noir» [1985], in *Die Anthropophorne Kamera*, Zurich : Chronos, 1997, p. 122.

bien que le «je» reste hors-champ (comme si le monde du film n'était pas encore pleinement «peuplé», se trouvait en attente de personnages). On comprend dès les premières phrases que cette voix nous parle depuis le point d'arrivée du récit, et qu'il s'agira de regarder en arrière («J'emballe mes effets dans le châle que ma mère portait au marché, et je quitte ma vallée. [...] Je laisse derrière moi cinquante ans de souvenirs»). L'audiovisualisation du monde qui surgit grâce à la réminiscence s'opère à travers deux procédés visuels: d'abord, la caméra s'avance en direction de la fenêtre et nous fait découvrir le paysage; ensuite, un fondu puis un noir déclenchent le changement d'époque. La «situation-cadre» est considérée ici de façon tout à fait littérale: c'est le surcadrage de la fenêtre «ouverte sur *un* monde» qui concrétise l'enchâssement du passé dans le présent. Le texte de la voix-over thématise explicitement la parenté entre passé (les souvenirs d'un personnage) et fiction (celle du film auquel ce prologue nous introduit):

«C'est étrange de penser que l'esprit oubliera tant de ce moment tout en gardant clairement le souvenir de ce qui s'est passé des années plus tôt, d'hommes et de femmes décédés longtemps auparavant. *Pourtant, qui peut dire ce qui est réel et ce qui ne l'est pas?* Puis-je croire au départ de mes amis quand leurs voix résonnent toujours splendidement à mes oreilles? Non, et je continuerai à dire non, car à mon esprit, ils sont toujours vivants. [...] Je ferme donc les yeux sur ma vallée d'aujourd'hui et elle disparaît. *Et je la vois comme elle était quand j'étais enfant.*»

La seconde phrase de cette citation, absente comme toutes celles de cet extrait du roman de Richard Llewellyn, peut paraître de prime abord quelque peu étrange. En effet, la motivation de cet énoncé n'est pas uniquement diégétique, mais relève de la fonction «contractuelle» du prélude: on nous signifie qu'il s'agira de *croire* à la réalité de cette représentation d'un passé fictionnel. Pour cela, on assigne au spectateur la place de «l'esprit» du narrateur, on lui confère un pouvoir ubiquitaire lui permettant de quitter, au gré du travelling, l'ici-maintenant de la pièce. Le noir apparaît à l'instant même où le narrateur prétend «fermer les yeux». Or le dernier temps présent de ce passage est significativement celui du verbe «voir» qui correspond à l'activité nécessairement «présente» du spectateur, alors que tout ce qui défile sera ensuite marqué du

sceau de l'antériorité («La mine ne faisait que commencer à noircir la verdure...»). Présentifier, c'est rendre «vivant». Ce texte par lequel s'actualise le passé du souvenir dans le présent de l'image souligne l'importance de la voix, instrument de récréation d'un monde révolu au service d'une réfutation de la mort. Non pas la voix-over, qui se manifeste sous la forme de la présence-absence, mais les voix des êtres disparus qui «résonnent toujours splendidement à [ses] oreilles». C'est avec l'apparition d'une voix *in*, celle de la sœur qui mentionne le nom du narrateur (le constituant par là en être de fiction), que nous entrons véritablement dans ce monde, car les discussions entre le jeune Huw et son père restent quant à elles muettes. Par ailleurs, le récit itératif qui nous est fait des occupations de la communauté de mineurs est accompagné d'une chanson, annoncée *over* comme étant l'opérateur de la transition de l'individuel au collectif («Quelqu'un entonnait une chanson, et la vallée s'emplissait du son de nombreuses voix»). Le début du film de Ford problématise véritablement les différents usages de la voix au cinéma: la voix-narration comme instigatrice d'un monde, la voix-action comme «animation» d'êtres décédés, et la voix-attraction du chant comme manifestation du collectif.

Même s'il n'est pas toujours introduit de façon aussi univoque que dans *Qu'elle était verte ma vallée*, le récit au passé constitue l'usage dominant de la voix-over au cinéma. Il est vrai que le texte *over* peut également être conjugué au présent, notamment dans le cas du monologue intérieur, mais il est par contre fort rare qu'il se rapporte au futur. Comme le note Châteauvert, «le récit proleptique est exceptionnel au cinéma, le passage au futur se révélant la plupart du temps une simple ellipse»<sup>622</sup>. Cette remarque s'explique selon moi par le fait que l'image a tendance à ancrer les assertions verbales relatives au futur dans le présent de l'action visualisée. Une tension se crée entre présent et futur qui, dans le cas du récit rétrospectif, n'est probablement suspendue qu'en raison des liens dont nous avons discuté entre passé et fiction. Il existe certes des films comprenant des inserts proleptiques (à l'exemple de *La Guerre est finie*, Alain Resnais, 1966), mais ils ne proposent que très rarement un enchâssement aussi complet qu'un *flash-back* audiovisuel. Alors que la prolepse verbale est happée par le présent de l'image, le passé de la voix-over tend inversement

622. J. Châteauvert, 1996, *Des mots...*, *op. cit.*, p. 171, note 19.

à marquer l'ensemble du monde audiovisuelisé du sceau de sa valeur prétéritive.

Dans les films non fictionnels, la voix-over endosse la même fonction intégrative que l'on trouve dans la fiction, mais le type d'organisation textuelle dominant n'y est pas forcément d'ordre narratif. En effet, les non-fictions possèdent généralement un degré de narrativité inférieur à celui qui caractérise le régime courant de la fiction. C'est pourquoi j'ai proposé de qualifier «voix-explication» cet autre versant du pôle de l'intégration. Fiction et non-fiction s'opposent comme deux pôles plutôt que comme deux catégories strictement exclusives. L'hybridité de certains films du point de vue de leur degré de fictionalité peut de ce fait produire un mélange de voix-narration et de voix-explication. La perception des temps passés varie alors au sein d'un même film.

#### VI.4 «CORPS» DU FILM, CORPS DU LOCUTEUR

Plutôt que de nous éloigner de l'univers diégétique, la voix-over, en dépit de sa position de retrait par rapport aux contingences de ce monde, nous y conduit. On objectera certes que l'atmosphère candide et presque onirique de *Qu'elle était verte ma vallée* fait figure d'exception, même au sein du cinéma hollywoodien de l'après-guerre. Il est vrai que d'autres films de cette époque suscitent une plus grande étrangeté qui tient souvent à l'indécision de l'ancrage du moment de l'énonciation verbale dans l'univers diégétique.

Cette caractéristique concerne notamment une partie de l'œuvre de Joseph L. Mankiewicz, dont *La Comtesse aux pieds nus* (*The Barefoot Contessa*, 1954). Dans la séquence d'ouverture de ce film, nous découvrons la situation-cadre de l'enterrement – «situation» bien plus que «récit», puisque cette scène est totalement dépourvue d'actions (et de sons synchrones qui vectoriseraient l'écoulement du temps), la cérémonie demeurant hors-champ alors que la caméra s'attarde sur le public immobile et sur la statue érigée en souvenir de la défunte actrice – qui donnera lieu à la visualisation des différents narrateurs. Outre le bruit d'ambiance de la pluie, seule une mélodie de guitare se fait entendre. Le spectateur l'associera rétroactivement, grâce à un insert sur un guitariste dans le club de Madrid, à l'époque du premier *flash-back*, ce qui place en fait ce prélude sous le signe du passé. Contrairement à *Qu'elle était*

*verte ma vallée* où le médiateur entre le monde du film et le spectateur se tient au seuil du hors-champ, la caméra s'avance ici en direction du personnage interprété par Humphrey Bogart et l'isole de la foule pendant que nous entendons cet acteur (au timbre de voix immédiatement identifiable) débiter un récit en voix-over. À l'écran, Bogart ne remue pas les lèvres; de plus, le texte est conjugué au passé, même lorsqu'il se rapporte avec exactitude à la situation que nous sommes en train de voir (ce qui exclut tout « monologue intérieur » simultané aux actions montrées). Alors que, dans *Qu'elle était verte ma vallée*, chaque événement passé a le parfum de la madeleine de Proust, *La Comtesse aux pieds nus* ne procède pas d'une opposition entre présent et passé, mais nous projette d'entrée de jeu dans un monde que les temps verbaux passés connotent, en interaction avec le paratexte (notamment la reconnaissance de l'acteur), comme étant fictionnel. Néanmoins, ces temps verbaux conservent au niveau diégétique une valeur pré-téritive en signifiant de façon très explicite que nous assistons déjà à la fin du récit. Cette façon de nous introduire dans le film en commençant par la fin de l'histoire est une stratégie narrative fréquente parmi les films en voix-over: au-delà de son enjeu thématique (voire philosophique) qui consiste à exprimer une vision déterministe de la condition humaine, ce principe vise également à accorder ostensiblement la préséance à la narration en exhibant la « détermination rétrograde » qui, selon J.-M. Adam, sous-tend toute organisation narrative<sup>623</sup>. Dans cette introduction de *La Comtesse aux pieds nus*, la déliaison vocale momentanée entre le personnage visualisé et le personnage-locuteur, ainsi que la distance temporelle (accrue par l'analepse enchâssée du cabaret madrilène) nous désolidarisent certes du personnage de l'écran, mais nous rapprochent de la voix qui nous apparaît d'autant plus « présente » que la situation est décrite au passé. La complicité avec le spectateur est d'ailleurs renforcée par le piquant des remarques du narrateur qui recourt au vocabulaire « métanarratif » des scénaristes et metteurs en scène de cinéma (il parle de « décor », de « mauvais films », de l'évolution « début-milieu-fin », « d'un fondu à l'autre », etc.). Cette dimension réflexive, omniprésente dans *La Comtesse aux pieds nus* mais accentuée ici, convient tout à fait au « seuil » du film où, dans le prolongement du générique, se rencontrent le monde du spectateur et

623. J.-M. Adam, *Le Récit*, op. cit., p. 87.



l'univers diégétique. Par ailleurs, ces expressions sont motivées diégétiquement par la profession du protagoniste (cinéaste), si bien qu'elles contribuent à construire ce personnage de fiction, défini par son professionnalisme et son caractère désabusé. Même singulièrement plaquée sur un visage immobile dont on attendrait qu'il s'exprime en voix synchrone, la voix-over se fait le véhicule de l'immersion diégétique, se présente comme un guide. La séquence des obsèques n'est d'ailleurs qu'une phase intermédiaire du voyage dans le temps (et dans la fiction) auquel nous invitent les voix-over de Bogart et d'autres narrateurs masculins qui participent à la représentation kaléidoscopique de Maria Vargas (Ava Gardner), figure de femme à la fois centrale et évanescence. L'énonciation polyphonique de ce film ainsi que la scène de la gifle (seul moment de « focalisation interne multiple » au sens de Genette) ne font que renforcer l'impression que la voix-over occupe une position « originelle » par rapport au monde décrit.

En dépit de sa contribution décisive à l'élaboration de la diégèse, la voix-over du prologue de *La Comtesse aux pieds nus* n'est pas aussi lisse et transparente que pourraient le laisser penser ces quelques remarques. En effet, il y a dans la coprésence de l'image d'un personnage impassible et de la voix-over de l'acteur qui incarne ce personnage un soupçon de l'« inquiétante étrangeté » qui caractérise la déliaison vocale. Pour rendre compte de cet effet sur le spectateur, il faut donc quitter le niveau de l'énonciation narrative pour s'intéresser à sa figuration textuelle. La « voix » est une catégorie qui, selon Genette, concerne à la fois les niveaux énonciatifs (question de l'« extradiégéticité ») et les rapports entre le narrateur du récit-cadre et les personnages du récit enchâssé (homo- ou hétérodiégéticité). Au cinéma, il est nécessaire d'ajouter la relation entre la voix-over et l'image en tenant compte de l'éventuelle visualisation du corps d'un personnage qui est aussi le narrateur. Cette relation intervient dans le cas d'un narrateur homodiégétique, c'est-à-dire lorsqu'il appartient lui-même, en tant que personnage, à l'histoire qu'il raconte.

Un film comme *Chaînes conjugales* (*A Letter to Three Wives*, Joseph L. Mankiewicz, 1949) met en évidence l'importance de la monstration filmique : selon la typologie de Genette, la narratrice extradiégétique Addie Ross peut être considérée comme homodiégétique. Or il faut souligner que ce personnage, bien qu'appartenant au niveau (intra)diégétique, n'apparaît jamais à l'écran sous

une forme individualisée. Éternelle absente dans les séquences en *flash-back* (même son portrait photographique contemplé par les personnages du film demeure hors-champ), elle n'est entraperçue qu'une seule fois, de dos. S'il fallait interpréter cette situation en des termes genettiens, on dirait qu'il s'agit d'une « paralipse » (« le récit ne saute pas, comme dans l'ellipse, par-dessus un moment, il passe à côté d'une donnée »)<sup>624</sup>. Toutefois, contrairement au texte littéraire où la question des modalités de visualisation ne se pose pas, il est nécessaire de prendre en compte au cinéma la façon dont le personnage apparaît à l'image ou en est évincé. On constate par exemple dans *Chaînes conjugales* que le pouvoir d'ubiquité et de manipulation d'Addie (interprétée vocalement par Celeste Holm, non créditée au générique) est accru par le caractère fantomatique de ses manifestations intradiégétiques. La question du « corps » du narrateur dans un médium qui permet son incarnation et joue avec sa désincarnation implique à mon avis de reconsidérer les catégories qui, selon Genette, relèvent de la « voix ».

La notion de « narrateur extradiégétique » demande par conséquent à être précisée lorsqu'on s'en sert pour décrire un film. Genette a conscience des difficultés que soulève la distinction entre « fait de niveau » et « fait de personne », puisqu'il juge utile d'y revenir dans *Nouveau discours du récit* en donnant l'exemple du roman de Lesage :

« Gil Blas est un narrateur extradiégétique parce qu'il n'est (*comme narrateur*) inclus dans aucune diégèse, mais directement de plain-pied, quoique fictif, avec le public (réel) extradiégétique; mais puisqu'il raconte sa propre histoire, il est en même temps un narrateur homodiégétique. »<sup>625</sup>

Au vu de la nécessaire consistance du monde créé par l'image ciné-photographique, il paraît difficile d'imaginer qu'un quelconque élément de la fiction du film puisse n'être « inclus dans aucune diégèse ». La diégèse est « toujours déjà là », posée dès la première image du film. L'unique façon d'échapper totalement au statut diégétique consisterait, à la rigueur, à ne pas connaître de visualisation et, en tant que phénomène acoustique, à ne pas être

624. G. Genette, *Figures III, op. cit.*, pp. 92-93.

625. *Ibid.*, p. 56.

perçu des personnages diégétiques, ce qui est le cas de certains narrateurs *over*, mais non de tous les narrateurs que Genette qualifierait d'«extradiégétiques». En effet, même si pour Genette le «fait de relation» ne concerne que l'opposition homodiégétique *vs.* hétérodiégétique, il faut bien convenir que le niveau extradiégétique n'existe que *comparativement* à un niveau intradiégétique. Durant la séquence de la villa genevoise du début de *La Fiancée de Frankenstein* (voir le chapitre V: Du monstre à sa fiancée, «création» frankensteinienne de la voix-narration), il est impossible de qualifier ce niveau d'«extradiégétique» avant que ne se soit opéré l'enchâssement du récit de la romancière. Cette appréhension rétrograde du niveau narratif représente le mode dominant de la lecture des récits filmiques à enchâssements. La définition que Genette donne ici de l'extradiégéticité pose en outre problème (et là réside vraisemblablement la cause de la fréquente «confusion» qu'il évoque) du fait qu'elle postule la coexistence de deux mondes, celui qu'occupe le narrateur et celui auquel son récit donne lieu (on se gardera par contre d'assimiler le niveau extradiégétique à celui des spectateurs réels du film, sous peine de confondre cinéma parlé et cinéma parlant). Or, dans le cas d'«analepses homodiégétiques», c'est-à-dire de *flash-backs* portant sur la même ligne d'action que le «récit premier»<sup>626</sup>, le monde du passé est identique à celui du présent, ce qui relativise grandement la position d'extradiégéticité du narrateur, quand bien même on rapproche l'instauration d'une fiction d'un «quasi-passé». La notion de «récit premier» est significative de la démarche du narratologue: il s'agit pour lui de poser dans un premier temps une instance de laquelle se distingue ce qui relève de sa production. Au cinéma, ladite instance ne peut selon moi jamais être située en dehors de la diégèse, à moins qu'il ne s'agisse d'une voix-*over* qui n'a jamais connu d'ancrage *in* (là aussi, l'appréhension peut être retardée), ou d'organisations profilmiques nettement distinctes du monde du film comme en proposent certains prologues «théâtraux». Daniel Percheron, dans l'un des premiers écrits d'inspiration narratologique portant sur le statut des voix au cinéma, notait déjà que le qualificatif «extradiégétique» posait problème lorsqu'il s'agissait de l'appliquer à des sons, et qu'il fallait entendre le préfixe «extra» «plus comme supplément que comme dehors»<sup>627</sup>. Ce constat l'amenait à proposer le néologisme

626. *Ibid.*, p. 92.

627. D. Percheron, «Le son au cinéma...», *art. cit.*, p. 82.

«paradiégétique», mais aussi (et c'est là qu'on ne peut plus le suivre aujourd'hui) à remettre en cause la pertinence de la notion de «diégèse» («la diégèse en effet, autonomisation abusive du narré, est une notion difficilement tenable»). En dépit du consensus qui règne chez les narratologues, le terme «extradiégétique» a donc soulevé passablement de problèmes.

Dans la séquence d'ouverture de *La Comtesse aux pieds nus*, la voix proférée par Bogart est bien «extradiégétique», et, corrélativement, le personnage qu'il interprète est intradiégétique. Par contre, si ce personnage était montré dans ce même environnement en voix synchrone comme un présentateur de prologue, on ne pourrait pas le considérer comme extradiégétique, fût-ce comparativement à son récit en *flash-back*, car la diégèse filmique, même réduite, s'est déjà mise en place. Ce problème découle en fait d'une confusion entre diégèse et histoire (voir *infra*). Mon intention n'est pas de décaler d'un niveau la partition de Genette (extradiégétique/intradiégétique/métadiégétique), car cette modification conduirait à multiplier dans certains films les préfixes «méta», mais plutôt de réserver l'appellation «extradiégétique» pour les seuls narrateurs totalement désincarnés (en tant que narrateur), et d'aborder l'enchâssement narratif en opposant simplement la situation-cadre (ou récit-cadre si elle s'inscrit dans une structure narrative) au récit enchâssé (en précisant le niveau s'il y en a plusieurs). Ainsi, on ne fait pas l'impasse sur la nécessaire «corporité» de tout univers filmique. Le terme «extradiégétique» est une dénomination pratique qui connaît néanmoins des limites: par opposition au boniment, on pourrait dire que toutes les voix d'un film parlant sont «diégétiques», y compris la voix-over. L'intégration de premier niveau rend en effet ces voix solidaires de la représentation filmique. Par ailleurs, en participant aux phénomènes de coréférence, la voix-over interagit avec d'autres canaux informatifs dont la source est explicitement ancrée dans le monde du film.

En tant que «voix narrative», la voix-over est associée à un facteur d'unicité. André Gaudreault utilise significativement, entre guillemets, le terme «univocité» pour qualifier le narrateur scriptural, terme dérivé de «voix» qui est loin d'être univoque<sup>628</sup>. Les théories narratologiques reposent sur une personnalisation qui leur

628. A. Gaudreault, *Du littéraire...*, *op. cit.*, p. 91.

permet de concevoir leur objet comme une totalité («l'œuvre» cohérente d'un «auteur», ou le «récit» comme entité close). Or, comme l'a noté Mary Ann Doane, cette unité est toujours imaginairement liée, dans un médium audiovisuel comme le cinéma, aux rapports entre le corps et la voix :

«Ainsi, même lorsqu'on utilise des sons asynchrones ou «sauvages», l'attribution fantasmatique d'une unité au corps n'est pas perdue. Elle est simplement déplacée – le corps *dans* le film devient le corps *du* film. Ses sens travaillent en tandem, car la combinaison du son et de l'image est décrite en termes de «totalité» et de composante «organique.»<sup>629</sup>

À titre d'extension du corps virtuel d'un personnage, la voix-over constitue également, de par la cohérence qu'elle confère à un ensemble fragmenté de plans, le «corps» du film. On peut par contre se demander ce qu'il advient de cette conception unitaire lorsqu'elle est confrontée à l'image du corps de celui qui la profère sans pour autant perdre, dans le passage concerné ou à d'autres moments du film, son statut *over*. Dans *La Comtesse aux pieds nus*, la rencontre entre la voix de Bogart-narrateur et le corps muet de Bogart-personnage a pour conséquence d'exhiber la déliaison fondamentale. C'est d'ailleurs peut-être pour conjurer l'effet de cette prise de conscience de la dissociation inhérente au dispositif que le texte *over* recourt si ostensiblement (mais avec désinvolture) à la réflexivité : les références participent du contrat de la lecture fictionnalisante tout en augmentant paradoxalement la crédulité du spectateur<sup>630</sup>. La fragmentation corporelle est en outre illustrée dès le début de l'enchâssement avec l'exhibition du corps de la danseuse dont le spectateur du film, contrairement aux spectateurs diégétiques ébahis, n'aperçoit que la main bord-cadre ainsi que l'ombre (ce hors-champ compense en quelque sorte la paralepse induite par l'absence du personnage «focal» interprété par Bogart). *La Comtesse aux pieds nus* repose sur l'impossibilité d'établir la cohérence d'un individu singulier comme Maria, ce qu'expriment le rapport initialement instauré entre voix et corps, puis la polyphonie de l'organisation énonciative de ce film-puzzle.

629. M. A. Doane, «The Voice in the Cinema...», *art. cit.*, p. 35.

630. Voir Ch. Metz, *Le Signifiant imaginaire*, *op. cit.*, p. 51.

Entre l'ancrage de la voix-narration dans un corps visualisé et son intégration au « corps » du texte filmique, différents cas sont possibles en fonction du mode d'apparition, à l'image, de celui qui peut être associé à un locuteur. C'est l'identification de l'appartenance d'un corps à la voix perçue qui entraîne chez l'audiospectateur la mise en relation de l'image et de la voix. Ce processus s'effectue soit de manière explicite par une visualisation synchrone préalable de l'acteur, soit de manière implicite grâce au savoir intertextuel du spectateur. En faisant débiter *La Comtesse aux pieds nus* par la voix-over de Bogart, Mankiewicz risquait peu qu'elle ne soit pas associée à l'acteur, puisqu'il s'agit d'une star de l'époque. La visualisation du corps du narrateur (et surtout de son visage) s'effectue de deux manières lorsqu'il y a enchâssement narratif : le film peut montrer soit le narrateur de la situation-cadre, soit un personnage de l'histoire qui est de surcroît le narrateur homodiégétique. Lorsqu'il n'y a pas d'enchâssement, la coprésence du corps de l'acteur (dont nous avons par ailleurs entendu la voix dans la diégèse) et d'une manifestation vocale *over* est également possible, car cette voix peut se présenter comme l'expression des « pensées » du personnage de l'écran. Dans de tels cas, le corps du locuteur tend à concurrencer le « corps » du film. Le degré de déliaison de la voix-over est proportionnel à l'incapacité physique dans laquelle le personnage visualisé se trouve de proférer les paroles *over*. On peut mentionner, dans un ordre croissant de déliaison et à titre de typologie non exhaustive, les cas de figure suivants, caractérisés par l'attitude du personnage visualisé pendant que l'on entend *over* l'acteur qui l'incarne :

- **Acousmatisation** : le locuteur est montré en voix synchrone avant de devenir *over*;
- **Indétermination de l'ancrage diégétique** : le visage de l'acteur reste hors-champ, ce qui crée une hésitation entre *off* et *over* (prologue d'*Hiroshima mon amour*) ; cas particulier de hors-champ : le visage d'un locuteur présent à l'écran demeure caché, par exemple dans l'ombre, comme dans le prologue du *Faux coupable* (*The Wrong Man*, Alfred Hitchcock, 1956) ;
- **Plaquage** : l'acteur ne remue pas les lèvres, mais manifeste au niveau de la mimique et de la gestuelle des signes qui sont en adéquation avec le contenu des paroles *over*, notamment grâce à une synchronisation vocogestuelle large. Situation caractéristique du

« monologue intérieur », par exemple dans *Duel* (Steven Spielberg, 1975) ;

– **Dissociation** : l'acteur ne remue pas les lèvres et paraît complètement indifférent au contenu des paroles *over* (ouverture de *La Comtesse aux pieds nus*) ; la dissociation, qui peut théoriquement être permanente, est généralement rétablie à la faveur d'une désaccousmatisation (qui peut s'effectuer, comme dans le film de Man-kiewicz, à un autre niveau narratif) ;

– **Doublage** : l'acteur remue les lèvres, mais aucun son *in* ne se fait entendre (*Le Roman d'un tricheur*) ;

– **Dédoublément** : l'acteur parle, sa voix audible se mêlant à la voix-*over* ; ce cas est peu fréquent étant donné la perte d'intelligibilité qu'il implique ; il entraîne généralement une mise en retrait de la voix *in* qui passe à l'arrière-plan sonore (comme lorsque Danny De Vito discute avec son banquier et commente simultanément la situation en voix-*over* dans *Casino*, Martin Scorsese, 1995) ; dans *Butterfly Murders* (Tsui Hark, 1979), la discussion entre le chef de guerre Tien Fung et le lettré Fong dans le souterrain du château provoque la superposition de la voix *in* et de la voix-*over* de ce dernier.

Plutôt que de s'en tenir au modèle abstrait de la « voix » au sens de Genette, il me semble donc intéressant de concevoir la voix-*over* dans les relations qu'elle entretient avec les personnages susceptibles d'être identifiés comme des locuteurs *over*. Les indécisions momentanées et les processus rétrospectifs sont symptomatiques de l'étendue des possibilités qu'offre le spectre du détachement de la voix par rapport à son ancrage dans l'image. De l'entre-deux entre *in* et *off*, résultat d'une propension inaboutie à l'ancrage dans l'image naît le potentiel d'étrangeté de certaines voix-*over*. En prenant en compte ces cas-limites, il est possible d'examiner les niveaux narratifs dans leur spécificité cinématographique, c'est-à-dire en considérant également la représentation visuelle de l'humain. L'unité corporelle, pleinement réalisée dans le standard du synchronisme, constitue le paradigme central de la perception audiovisuelle considérée du point de vue de la voix.

J'ai tenté de montrer dans cette section certaines parentés entre ce qui est traditionnellement dit dans les études littéraires de la « voix narrative » et la façon dont on a appréhendé le fonctionnement de la voix-*over* au cinéma. Les théoriciens du cinéma ont en effet capitalisé sur des acquis issus d'autres champs tels que la poétique, la

narratologie poststructuraliste, la critique littéraire, etc., partageant avec ces disciplines une interrogation commune: c'est la fameuse question du « Qui parle? », posée à partir des travaux de Roland Barthes<sup>631</sup>. L'année même où le texte de Barthes est publié, Christian Metz déclarait à propos du Narratif au cinéma: « L'impression que *quelqu'un parle* n'est pas liée à l'existence empirique d'un raconteur précis et connu ou connaissable, mais à la perception immédiate, par le consommateur du récit, de la nature langagière de l'objet qu'il est en train de consommer: parce que ça parle, il faut bien que *quelqu'un parle*. »<sup>632</sup> Le constat metzien s'inscrit dans des préoccupations théoriques centrales, dont on a vu en discutant les thèses de Laffay qu'elles sont passablement innervées par une réflexion sur la voix-over. Il m'a paru instructif d'observer combien le terme de « voix » joue un rôle fédérateur en dépit (ou justement en raison) de sa forte polysémie. En 1966, Metz affirmait au gré d'une réflexion sur les voix *in* et *off* dans certains films de la « modernité » qu'« il y aurait tout une étude à faire sur la *voix* chez un Godard ou un Resnais: c'est le problème du « Qui parle? »<sup>633</sup>. Le recours à l'italique est révélateur de la pertinence de cette composante du discours filmique d'un point de vue narratologique: le sémiologue fait simultanément référence au sens littéral du terme tout en ouvrant la voie à la typologie genettienne.

Les convergences que l'on observe autour de cette question de l'origine discursive s'expliquent peut-être par le modèle qui est sous-jacent aux théories des différents champs concernés. On peut le faire remonter à l'opposition platonicienne entre le récit simple (*haplè diègèsis*) et l'imitation (*mimèsis*)<sup>634</sup> où, dans la seconde, « le poète parle à la place d'un autre ». Cette représentation vococentriste de la narration se perpétue même lorsque le récit est exclusivement composé d'images produites par une machine. Le rapprochement est bien sûr plus aisé dès lors que le film présente du verbal et obéit à une structure narrative. C'est pourquoi on a pu constater combien l'étude de la voix-over a constitué un objet privilégié des approches d'obédience énonciative au cinéma. Pour

631. Roland Barthes, « Introduction à l'analyse structurale des récits », *Communications*, N° 8, 1966, repris in *Poétique du récit*, Paris: Seuil, 1977, pp. 39-40.

632. Christian Metz, *Essais sur la signification au cinéma*, tome 1, Paris: Klincksieck, 1966, p. 29.

633. Ch. Metz, *Essais sur la signification au cinéma*, tome 2, *op. cit.*, p. 215.

634. J'utilise ici la dichotomie commune. Pour les détails philologiques, je renvoie à A. Gaudreault, *Du littéraire au filmique*, *op. cit.*



Francis Vanoye par exemple, le « Qui parle ? » ne peut se poser relativement au récit filmique que si ce dernier est également « pris en charge par une (ou plusieurs) voix »<sup>635</sup>, c'est-à-dire à la seule condition qu'il comprenne une voix-over – il illustre significativement son propos en mentionnant des films comme *Jules et Jim* ou *La Comtesse aux pieds nus*. Les glissements entre voix-over et voix-narrative sont constants, ainsi que le suggère Gaudreault dans la formule « la « voix » narrative du narrateur en « voix » off »<sup>636</sup>. L'usage des guillemets souligne toute l'ambiguïté de la frontière entre les deux notions.

Inversement, on a vu que Jean Mottet définissait le mode spectaculaire du vaudeville en opposant le caractère fragmentaire des actes à l'unicité de la voix narrative (voir chapitre V.3, L'attraction vocale de type « vaudeville »). On retrouve là cette bipolarité que j'ai proposé d'établir entre voix-attraction et voix-narration sur la base du critère de l'« intégration ». Les hypothèses théoriques relatives à la voix narrative permettent d'expliquer le pouvoir intégrateur de la voix-over qui, en tant qu'indice évident de l'organisation discursive du film, se voit associée (par les théoriciens, mais aussi par les spectateurs) à l'instance ultime de production du film (au « méga-narrateur filmique » selon Gaudreault). Je discuterai ci-dessous plus spécifiquement la place qu'il convient d'accorder à la voix-over dans la structure de l'énonciation filmique. On peut déjà noter que cette question de la *perception* de la voix-over comme étant (illusoirement) un équivalent de la « voix narrative » doit être également prise en compte, tant en termes de stratégie narrative que de réception spectatorielle.

Ce que j'ai dit à propos du bonimenteur comme instance d'actualisation d'un discours verbal implicitement inscrit dans l'agencement des images vaut également pour la voix-over, même si les modalités de sa médiation diffèrent en raison du « premier degré d'intégration » qu'elle implique. Plus que toute autre manifestation vocale, la voix-over affiche son origine dans l'espace de la production. Néanmoins, elle lance également un pont vers l'activité cognitive du spectateur à laquelle elle peut paraître se substituer, alors que, souvent, elle l'intensifie en augmentant le nombre d'informations transmises. Le « discours intérieur » ne consiste pas

635. Francis Vanoye, *Récit écrit, récit filmique*, Paris: Nathan, 1989, p. 153.

636. A. Gaudreault, *Du littéraire... , op. cit.*, p. 143.

seulement, comme on l'a vu, à transposer le sens des images en équivalents linguistiques, mais aussi à organiser la structure syntagmatique du film sur le modèle du langage verbal. Ainsi, Tom Gunning a dit du montage effectué par Griffith à partir de 1908, basé sur un « système du narrateur », qu'il tendait à prendre la forme « d'une sorte de conférencier filmique intériorisé »<sup>637</sup>. En tant que verbalisation du discours intérieur, la voix-over, facteur de mise en forme du discours qui n'est pas totalement de nature linguistique (elle se situe dans les « limbes » entre verbal et « pensée »), se prête fort bien à une comparaison avec cette notion que les théoriciens de la littérature ont nommée « voix narrative ».

Il faut toutefois préciser que la voix narrative d'un film ne se résume pas à cette composante verbale, ainsi qu'en témoignent le statut facultatif de la voix-over et la lisibilité de nombreux films narratifs à voix-over lorsqu'on supprime cette dernière. En effet, si la voix-over « actualise » le discours élaboré par le montage du film (ou les connotations induites par d'autres pratiques filmiques), elle ajoute souvent peu de chose du point de vue narratif. Par exemple, la version « *director's cut* » (sans voix-over) de *Blade Runner* (Ridley Scott, 1982), qui a aujourd'hui remplacé la version originale dans laquelle Harrison Ford était un narrateur *over*, apparaît peut-être plus contemplative et moins liée au genre du « film noir » que la première, mais elle n'en reste pas moins compréhensible au niveau narratif. En fait, l'association de la voix-over à la voix narrative par l'audiospectateur est d'autant plus forte que la subordination de l'image au verbal est importante. Il en va ainsi lorsqu'un enchâssement est construit pour donner l'impression qu'il a été créé par la voix-over, ou lorsqu'une série d'images n'obéit pas intrinsèquement à une organisation narrative.

## **VI. 5 VOIX-OVER ET IMAGE : DEUX INSTRUMENTS POUR CONSTRUIRE UN MONDE**

Jean Châteauvert décrit l'interaction entre la voix-over et l'image en disant que le « mécanisme de référence permet au discours du narrateur de prendre appui sur le monde audiovisuel et crée l'impression que le matériel audiovisuel prolonge le récit verbal »<sup>638</sup>. Sa conception est informée par la volonté d'établir une adéquation

637. T. Gunning, *D. W. Griffith and the Origine...*, *op. cit.*, p. 93.

638. J. Châteauvert, *Des mots...*, *op. cit.*, p. 59.

entre le contenu des paroles et celui des images qui repose à mon sens sur la notion de «coréférence». Toutefois, ce qui est (généralement) vrai du discours verbal dans son ensemble ne l'est pas de toutes les unités qui le composent: nombre de lexèmes ne trouvent pas de référent dans le monde (audio)visualisé, ce qui ne les empêche pas de contribuer à construire ce monde. Car la référence en tant que telle, qu'elle soit verbale, iconique ou verbo-iconique, est destinée à la mise en place d'un monde (fictionnel ou renvoyant à la réalité) susceptible d'être communiqué. La dimension référentielle du discours nous conduit donc à examiner la célèbre notion de *diégèse* dont l'usage s'est généralisé avec la narratologie genettienne<sup>639</sup> après qu'elle a été réactualisée (sur la base de la «*diegesis*» platonicienne puis aristotélicienne) et précisée à la fin des années 40 dans le champ du cinéma par le filmologue Étienne Souriau<sup>640</sup>. Ce n'est probablement pas un hasard si le cinéma est précisément le premier moyen d'expression à avoir suscité ces réflexions sur le «monde» créé par un récit (car c'est peu dire que la narratologie littéraire a fourni plus de concepts à l'étude du cinéma qu'elle ne lui en a emprunté). L'apparente proximité de la représentation iconique avec la réalité, qui ne concerne pas la littérature d'un point de vue perceptif, appelait impérativement une distinction des niveaux telle que l'a entreprise Souriau. La hiérarchisation qu'il a proposée comprenait le «diégétique» dans une relation d'opposition à d'autres niveaux dont les dénominations connurent, à l'exception du «profilmique» (ce qui est placé devant la caméra au tournage), une moins grande postérité. Souriau donne la définition suivante du terme «diégèse»:

«Tout ce qui appartient, <dans l'intelligibilité> (comme dit M. Cohen-Seat) à l'histoire racontée, au monde supposé ou proposé par la fiction du film.»<sup>641</sup>

639. Notons que, dans *Figures III*, Genette ne fait pas mention de Souriau (il n'apparaît même pas dans sa bibliographie!), alors qu'il se réfèrera à lui dix ans plus tard en revenant sur son ouvrage, disant notamment qu'il «dérive toujours (comme Souriau, bien sûr) *diégétique* de *diégèse*, jamais de *diégésis*» (*Nouveau discours du récit, op. cit.*, p. 13).

640. La filmologie fut un courant d'études sur le cinéma qui reposait sur le principe de l'interdisciplinarité. Son essor a grandement contribué, en France, à la prise en compte du médium dans le domaine académique. Initié en 1946, ce mouvement disparaît au début des années 60. Sur l'implantation institutionnelle de la filmologie, voir André Chaperon, «L'institut de filmologie, une tentative d'interdisciplinarité», *Études de Lettres*, N° 2, 1993.

641. Étienne Souriau, «L'impression de réalité au cinéma: les phénomènes de croyance», in E. Souriau (dir.), *L'Univers filmique*, Paris: Flammarion, 1953, p. 7. Voir également la rubrique «Diégèse» in Anne Souriau (dir.), *Vocabulaire d'esthétique*, Paris: PUF, 1990, pp. 581-583. Étonnamment, Anne Souriau s'attribue la paternité de cette notion, qu'elle aurait proposée en 1950.

Il me paraît utile de considérer chacun des termes de cette définition qui, lorsqu'elle est tronquée par les continuateurs de Souriau, tend à réduire cette notion à celle d'«histoire» au sens de Genette. Ainsi ce dernier parle-t-il des «relations temporelles entre récit et diégèse», ce qui postule une stricte équivalence entre «histoire» et «diégèse»; le narratologue se corrige d'ailleurs dans *Nouveau discours du récit* en disant de la diégèse qu'elle «est bien un univers plutôt qu'un enchaînement d'actions (histoire)»<sup>642</sup>. Cette correction ultérieure est révélatrice de la confusion qui a traditionnellement régné dans les discussions portant sur cette notion.

Revenons à Souriau: tout d'abord, l'expression «dans l'intelligibilité», que le filmologue emprunte à son collègue Gilbert Cohen-Séat, signifie d'entrée de jeu que la diégèse est le produit d'un travail cognitif du spectateur; cette précision est d'importance puisque, à la suite de théoriciens américains du cinéma comme David Bordwell, il me paraît légitime d'opter pour une approche de la réception filmique en termes d'«inférences» spectatorielle<sup>643</sup>. Je qualifierai donc d'«inférences diégétiques» les déductions effectuées par le spectateur lorsque, sur la base d'informants livrés par le film, il «complète» mentalement ce qui n'est présent qu'implicitement dans le film. Ce type d'inférences est rarement convoqué de façon autonome: c'est également en comparant le *monde du film* avec celui de son expérience quotidienne (inférences encyclopédiques) ou avec d'autres mondes imaginaires (inférences intertextuelles) que le spectateur actualise un grand nombre de (pré)suppositions. Le terme de «monde» utilisé par Souriau me paraît important pour la définition de ce niveau de «réalité» filmique: le filmologue signifie par là qu'il s'agit d'une catégorie sémantique, d'un ensemble de référents, et non d'un mode d'organisation du discours comme l'est le récit. Enfin, en accord avec la référence liminaire à l'activité cognitive du spectateur, Souriau note que ce monde ne se limite pas à ce que montre le film, mais comprend également ce qui, sur cette base, peut être *supposé* par le spectateur. La notion de «diégèse» présente l'avantage de ne pas confiner l'entreprise analytique à l'immanence du texte, mais de soulever des interrogations relatives à la façon dont le spectateur se représente le monde proposé par le film.

642. Voir G. Genette, *Figures III*, *op. cit.*, p. 75, et *Nouveau discours du récit*, *op. cit.*, p. 13.

643. Anne Souriau précise à cet égard que «les éléments diégétiques non directement manifestés dans l'œuvre peuvent en être déduits par le raisonnement [...]» (*Vocabulaire d'esthétique*, *op. cit.*, p. 582).

Il nous faut cependant émettre une réserve quant à cette définition : à mon sens, le terme de « fiction » utilisé par Souriaut ne doit pas conduire à restreindre l'application de la notion de « diégèse » aux seuls films fictionnels. On comprend certes, en cette période de « réalisme ontologique » bazinien, la volonté, manifeste chez le filmologue, de pointer toutes les « irréalités » de la représentation, mais, comme je l'ai montré dans *La Fiction au cinéma*, cette « construction » qu'est la diégèse vaut tout autant pour les documentaires (et donc également pour la voix-explication). Quel que soit le type discursif dominant au sein du texte *over*, la fonction « intégrative » de ce dernier assure une optimisation de la cohérence et de la clôture diégétique. Un film non fictionnel demande, au même titre qu'une fiction, que le spectateur « suppose » certains éléments de son monde, par exemple lorsqu'ils se trouvent hors-champ. Cette construction mentale est indépendante du degré de fictionalité de ce qui est représenté. Il faut en effet préciser qu'une lecture n'est pas « fictionalisante » du seul fait que le spectateur recourt à son imagination. Du point de vue de l'élaboration diégétique, on peut dire qu'un seul critère, relevé par André Gardies, distingue la fiction du pôle opposé (situé sur une échelle graduée) : l'*autonomie* du monde du film par rapport au monde réel<sup>644</sup>. La fiction *produit* un monde, le documentaire *restitue* le réel. Cette relation au monde factuel ne signifie pas pour autant que les films non fictionnels se dispensent d'organiser cette restitution, ni, à l'inverse, que les fictions sont construites indépendamment de tous liens avec le monde de référence du spectateur (et de ceux qui les élaborent).

La diégèse ne ressortit donc pas en propre au domaine de la fiction. On peut en outre se demander si elle appartient nécessairement à une « histoire racontée ». En effet, la pertinence d'une distinction théorique entre les données diégétiques et narratives me semble nécessaire. La parenté que ces deux catégories entretiennent n'invalide pas la nécessité de ce distinguo mais, au contraire, la renforce. On pourrait dire que cette distinction recoupe celle que posait Genette entre description et narration dans *Figures II*, où il soulignait à la fois la préséance de la seconde et l'interdépendance des deux régimes.

644. A. Gardies, *Le Récit filmique*, Paris : Hachette, 1993, p. 43.

«La description pourrait se concevoir indépendamment de la narration, mais en fait on ne la trouve pour ainsi dire jamais à l'état libre; la narration, elle, ne peut exister sans description, mais cette dépendance ne l'empêche pas de jouer constamment le premier rôle.»<sup>645</sup>

Il ne s'agit pas d'assimiler diégèse et description, car ce choix ne ferait que déplacer la confusion entre diégèse et récit vers un autre type discursif: de nombreux théoriciens – dont Adam et Petitjean en analyse des discours, ou André Gardies pour le cinéma – ont montré que la «description» consistait en une procédure de mise en texte bien spécifique. La diégèse est en revanche un ensemble sémantique composé d'une série de référents virtuellement exploitables, que cela soit pour décrire, raconter, argumenter, etc. Toutefois, la façon dont Genette décrit la relation de la description au récit me paraît s'appliquer aux liens qui unissent ce dernier à la diégèse: là aussi, une indépendance des deux notions est conçue sur le plan théorique, mais l'analyste est ensuite confronté, dans les textes, à une relation mutuelle. D'ailleurs, l'idée communément associée à la notion de «description», basée sur le paradigme/état *vs.* action/, convient à la diégèse. Significativement, Genette explique le caractère premier de la description en émettant l'hypothèse selon laquelle «les objets peuvent exister sans mouvement, mais non le mouvement sans les objets»<sup>646</sup>. Il est évident que cette opposition du statisme et du dynamisme est d'autant plus prégnante dans le cas de l'image animée: parmi les innombrables éléments que nous apercevons à l'arrière-plan de l'image, il se peut qu'aucun n'intervienne à proprement parler dans le récit filmique, bien que tous eussent pu le faire.

Afin de clarifier ce propos, on peut faire la comparaison suivante: dans le bac à sable où l'enfant s'invente une histoire et dont le carré découpe une certaine portion du monde réel, tout ressortit à la diégèse (les différents jouets, voitures et figurines, ainsi que le sable), à l'exception de l'enfant lui-même (il est en quelque sorte «extradiégétique», et se fait souvent énonciateur verbal à travers le «discours égocentrique» étudié par Piaget). À partir de ces éléments, il peut construire des histoires en animant véhicules et personnages, qui,

645. G. Genette, «Frontières du récit», *art. cit.*, p. 57.

646. *Idem.*

toutes, seront déterminées par les objets mis initialement à sa disposition, à moins qu'il ne recoure à son imagination pour se représenter mentalement ce qui lui manque. Au-delà du bac, bien sûr, le continuum de la réalité se poursuit, mais il n'est pas pertinent, car seule la parcelle de sable a été retenue pour constituer le «cadre» de l'histoire. Cette comparaison entre diégèse et activité ludique évoque la fiction, alors que la question des «données» mises à la disposition de l'enfant renvoie aux contraintes du documentaire. L'exemple du bac à sable éclaire les liens entre récit et diégèse: en fonction du référent que l'enfant assigne aux jouets disponibles, ceux-ci déterminent un certain nombre de «programmes» narratifs (une sorte d'horizon des possibles); en retour, l'actualisation d'un contenu narratif donné suscitera des besoins en supports réels ou imaginés.

Dominique Chateau a très bien décrit ce double mouvement entre diégèse et récit où la cohérence diégétique (souvent solidaire de la vraisemblance) est évaluée par rapport au récit, alors que celui-ci se voit à son tour sans cesse redéfini par les données diégétiques<sup>647</sup>. Le récit assure ainsi simultanément la conservation et l'expansion du monde diégétique, principes sur lesquels reposent les relations entre diégèse et récit. Michel Colin précise à propos de ces phénomènes de compréhension que «ce qui permet la construction progressive de la diégèse, c'est notamment l'interprétation des relations d'anaphore et de coréférence»<sup>648</sup>. Si ces deux mises en relation s'effectuent sur le plan syntagmatique de la bande-image, elles sont également à l'œuvre entre le verbal et l'image. Il est par conséquent nécessaire, si l'on veut saisir le rôle de la parole au cinéma, de prendre en compte la participation du signe linguistique à l'élaboration de la diégèse par des phénomènes de référence et de coréférence, tant sur un plan syntagmatique (le verbal construit le monde à partir d'informations verbales transmises antérieurement) que vertical (grâce à la coréférence verbo-iconique, elle-même susceptible de se déployer horizontalement selon les principes de la synchronisation). C'est pourquoi l'affirmation d'André Gardies selon laquelle, dans le documentaire, le monde diégétique «procède davantage par moments, par fragments, entrecoupé ou régi qu'il est par le commentaire verbal» me

647. D. Chateau, «Diégèse et énonciation», *Communications*, N° 38, 1983, p. 128.

648. M. Colin, *Langue...*, *op. cit.*, p. 22.

semble dépourvue de sens, car elle suppose l'absence momentanée du monde du film, alors que le spectateur l'a sans cesse à l'esprit<sup>649</sup>. Sous peine de priver la notion de «diégèse» de toute pertinence, on ne peut imaginer que le langage verbal «entre-coupe» le monde du film : il y a là un présupposé qui est probablement lié à la confusion induite par la notion d'«extradiégéticité» (voir *infra*). Bien au contraire : ne pouvant s'abstenir de référer à quelque chose, le verbal est nécessairement un facteur de construction diégétique (à l'exception peut-être de certains usages de la voix-attraction totalement dominés par les fonctions phatique ou conative). On a d'ailleurs vu que, sous sa forme orale, le verbal endosse généralement dans les films le rôle d'instigateur du monde audiovisuel, et cela d'autant plus lorsque sa source n'appartient pas à l'univers diégétique. En convoquant un mode de transmission premier et archaïque (la «performance» selon Zumthor), la voix-over, héritière entre autres pratiques des prologues *live*, suggère par sa seule présence matérielle (le grain, le timbre, l'accent, etc.) qu'une instance humaine est à la source du monde qui nous est donné à voir et à entendre. Elle se fait plus ou moins ostensiblement le véhicule premier de la représentation en fonction de la *relation de subordination* qu'elle instaure avec l'image. La soumission de l'image au verbal est proportionnelle au nombre de référents verbaux autonomes, ou du moins à la quantité de référents linguistiques dégagés de la coréférence verbo-iconique simultanée.

La façon dont Dominique Chateau envisage dans l'extrait cité ci-dessus l'implication du spectateur dans la diégèse exige en outre de considérer l'activité du narrataire comme un processus dynamique. Chateau isole deux composantes : la croyance et la construction de l'univers diégétique. Elles sont interdépendantes puisque le spectateur accepte les relations configurées par la diégèse s'il agrée aux lois proposées. Son assentiment est lui-même déterminé par des inférences intertextuelles (les genres filmiques, la connaissance du roman dont le film est l'adaptation, le «monde» associé à un «auteur», à une maison de production, à un acteur, etc.), narratives (tout fantaisiste que soit ce monde, il doit présenter, afin d'être communicable, une cohérence interne), encyclopédiques (il repose nécessairement en grande partie sur le monde de l'expérience quotidienne propre à la culture dont il est

649. A. Gardies, *Le Récit filmique*, op. cit., p. 43.



le produit), etc. En ce sens, la diégèse est à la fois le lieu d'exercice et le résultat de ce qu'Umberto Eco nomme, dans *Lector in fabula*, la «coopération textuelle» du lecteur.

La référence à Eco n'est pas fortuite puisqu'on peut voir des similitudes entre la notion de Souriau et certaines théories plus récentes menées dans le champ des études littéraires concernant les «mondes possibles», concept transposé de la théodicée leibnizienne dans le contexte de la logique<sup>650</sup>. Ce courant théorique a selon moi l'avantage, concurrentement aux études de l'«énonciation», de sortir définitivement l'analyse des récits de l'ornière structuraliste. Le récit n'est plus seulement un texte considéré dans sa dimension syntaxique, mais une organisation qui interagit avec l'élaboration progressive d'un monde dans l'esprit du narrataire. De nombreux commentateurs, dont notamment Umberto Eco, ont noté une certaine inadéquation entre la notion de «mondes possibles» telle que conçue par les logiciens et l'usage qu'en font les théoriciens de la littérature<sup>651</sup>. Ces différences tiennent surtout au fait que la création littéraire ne se réalise pas à l'intérieur d'un cadre contraignant initial (même si, dans une moindre mesure, il existe là aussi certaines contraintes, par exemple génériques) et que son monde est incomplet. En effet, le «monde possible» de la logique modale correspond à un ensemble maximal de propositions, alors que les fictions ne nous livrent qu'un nombre restreint d'informations sur les «états de choses» qui les composent, généralement en faisant primer la pertinence narrative. Sans entrer dans les détails de cette polémique, on peut remarquer que l'application de cette approche au récit littéraire a l'avantage d'avoir conduit à une représentation différente de la fiction. Au lieu d'être considérée comme un simulacre du réel, elle est abordée de façon positive, supplantant une tradition linguistico-logique qui a consisté à définir la fiction par son caractère «non sérieux». Dans cette optique «positive» qu'Umberto Eco qualifie de «constructiviste»<sup>652</sup>, un «monde possible» se définit comme un ensemble d'individus dotés de propriétés, soumis à des lois spécifiques et interagissant avec un environnement spatio-temporel donné.

650. Voir Christine Montalbetti, *La Fiction*, Paris: Flammarion, 2001, p. 236.

651. Umberto Eco, *Les Limites de l'interprétation*, Paris: Grasset, 1992, pp. 212-213. Voir la synthèse proposée par Jean-Marie Schaeffer, *Pourquoi la fiction?*, Paris: Seuil, 1999, pp. 206-207.

652. U. Eco, *Lector...*, *op. cit.*, p. 169.

Il est évident que le texte *over*, grâce à sa capacité de dénomination et de qualification, contribue largement à la construction du monde filmique. On peut considérer la diégèse comme l'équivalent d'un « monde possible », au sens où Eco propose d'entendre cette notion. Cette démarche permet à mon sens de renouveler, à partir des modèles théoriques considérés, la façon dont on a pensé le « récit filmique ». En effet, à l'instar de la « diégèse » selon la définition de Souriau, les « mondes possibles » incluent ce qui est *supposé* par le spectateur :

« Étant une construction culturelle, un monde possible ne peut être identifié à la *manifestation linéaire du texte* qui le décrit. Le texte qui décrit cet état ou cours d'événements est une stratégie linguistique destinée à déclencher une interprétation de la part du Lecteur Modèle. »<sup>653</sup>

La diégèse n'offre pas une « manifestation linéaire », réglée au cinéma par la synchronisation des phénomènes (co)référentiels, mais inclut l'ensemble des états de choses proposés par l'intégralité du film. Certes, on peut repérer comment la transmission progressive d'informants diégétiques est répartie sur l'axe temporel, et, réciproquement, ce que cette distribution implique en termes de participation du spectateur, mais la diégèse ne peut être envisagée comme un ensemble qu'une fois le film terminé. En outre, il n'y a pas d'ontologie de la diégèse : comme je l'ai déjà souligné, elle n'existe nulle part ailleurs que dans l'esprit du spectateur. Le film, loin de construire une diégèse de A à Z, s'en remet au spectateur chez qui il « déclenche » la *compréhension* (dans les deux sens du terme : on « comprend » ce qui est « inclus »). Sauf indications contraires, ce dernier inférera par exemple que tout ce qui caractérise le « monde actuel » de son expérience vaut pour le monde du film. Cette comparaison est soumise au « principe de l'écart minimal » dont j'ai discuté ailleurs les applications qu'elle pouvait connaître dans la détermination du degré de fictionalité d'un film<sup>654</sup>. Marie-Laure Ryan définit cette notion en ces termes : « Ce principe postule qu'à quelque moment que nous interprétions un message concernant un monde alternatif, nous reconstruisons ce

653. U. Eco, *Les Limites...*, *op. cit.*, p. 214.

654. A. Boillot, *La Fiction au cinéma*, *op. cit.*, pp. 141-147.

monde de sorte qu'il soit le plus proche possible de la réalité que nous connaissons.»<sup>655</sup> La dimension pragmatique de cette théorie est soulignée par Thomas Pavel lorsqu'il évoque les liens qui s'instaurent entre les mondes possibles et les deux pôles de la communication, soulignant le fait que «les écrivains et les lecteurs fictifs émergent, par simulation, des membres réels de la communauté culturelle, et conservent, dans le système fictionnel, la plupart des traits culturels ou biologiques des participants réels à l'échange»<sup>656</sup>. L'un des «traits biologiques» majeurs du monde d'expérience de l'audiospectateur dont les films de fiction parlants «héritent» est bien sûr celui de la voix: c'est en elle qu'il *se reconnaît* en tant qu'humain. Même lorsque le monde du film se situe aux antipodes de notre monde actuel, comme cela est le cas dans beaucoup de dessins animés (tant au niveau figuratif qu'en ce qui concerne les lois physiques qui le régissent), la voix est le lieu d'une rencontre, à moins que ne s'y marque une déliaison qui renvoie à son statut machinique. À l'instar du mouvement (quant à lui lié à l'«animé», et non spécifiquement à l'«âme»), la voix est un vecteur fondamental de cette interpénétration des mondes réel et fictionnel qui assure l'immersion fictionnelle.

La comparaison entre différents mondes alternatifs a mené les théoriciens à prendre en compte les *conditions d'accessibilité* existant entre ces mondes. Dans son analyse des stratégies dont use un texte littéraire pour nous introduire à son monde, Svetlana Vogeleer distingue deux processus sémantiques, la relation perceptuelle du voir et la relation épistémique du savoir, qui ne sont pas sans rappeler le couple notionnel ocularisation/focalisation que François Jost a dégagé de son étude sur le «point de vue» au cinéma, et qu'il proposait d'appliquer en retour au texte littéraire<sup>657</sup>. Ces deux modalités auxquelles recourt le narrateur scriptural selon Vogeleer doivent être complétées, au cinéma, par le caractère effectif de l'acte de perception – un roman crée la relation «X voit Y», alors qu'un film visualise Y – et par la dimension sonore (dans ce même cadre, Jost parle d'«auricularisation»), tant en ce qui concerne l'interaction avec le spectateur qu'au niveau des relations créées par

655. Marie-Laure Ryan, «Fiction, Non-Factuals, and the Principle of Minimal Departure», *Poetics*, N° 9, 1980, p. 403.

656. Thomas Pavel, *Univers de la fiction*, Paris: Seuil, 1986, p. 113.

657. Svetlana Vogeleer, «La relation point de vue et son application aux phrases existentielles initiales», in Walter De Mulder, Franc Schuerewegen et Liliane Tasmowski (dir.), *Énonciation et parti pris*, Amsterdam; Atlanta: Rodopi, 1992.

le texte. Car la voix n'est pas qu'un stimulus sonore, mais le support d'une parole qui, elle aussi, contribue à la construction du monde, convoquant ainsi simultanément, à un niveau textuel, la «relation du point de vue» examinée par Vogeleer. Un film peut nous faire entendre la voix d'un personnage, ou faire dire à la voix-over qu'un personnage parle ou qu'un autre l'entend (et même, comme dans la séquence enchâssée du shaman dans *Rashômon*, faire dire à la voix-over ce que les personnages visualisés sont censés dire). La voix-over constitue donc un facteur d'accessibilité au monde du film, tant sur le plan de la perception par l'auditeur que sur celui de la mise en scène textuelle de situations perceptives. En ce qui concerne le rapport de la voix-over au monde du film, il est donc indispensable de dissocier ces deux niveaux en postulant que la diégétisation s'effectue d'une part grâce à la présence même d'une manifestation vocale – l'audiospectateur est incité à inférer la «position» de l'énonciateur verbal fictionnel relativement à la diégèse, d'autre part à travers l'énoncé verbal produit par cet énonciateur. Le statut ambigu de cette voix qui appartient au film sans forcément être ancrée dans un monde la prédispose à l'exploitation du potentiel «diégétisant» des énoncés qu'elle profère. Lorsqu'elle joue le rôle de ce que l'on appelle significativement un *pont* sonore entre une situation-cadre et une situation enchâssée, la voix-over se fait le vecteur explicite de l'accession au monde du film. Ce cas de figure, très fréquent, laisse présager la pertinence de la théorie des mondes possibles pour rendre compte de la contribution de la voix-over aux processus d'intellection qu'engagent certains films.

Il n'est donc pas étonnant que Jean Châteauevert ait recouru précisément à cette théorie pour traiter de la voix-over. La question des «mondes possibles» ne guide cependant pas sa réflexion, mais vient s'inscrire dans le modèle premier de la responsabilité énonciative du locuteur: «Les individus et le cours des événements d'un monde possible ne sont pas actuels, mais relatifs à une attitude propositionnelle précise: *quelqu'un* affirme, croit, rêve, désire, etc., l'ensemble des événements et individus d'une diégèse»<sup>658</sup>. Dans cette optique, il s'agit d'analyser les modalités selon lesquelles le monde du film «dépend» de l'instance vocale *over* qui instaure une relation d'accessibilité à celui-ci (rêver, se souvenir, imaginer, mentir,

658. J. Châteauevert, *Des mots...*, *op. cit.*, p. 68 (je souligne).

etc.). Afin de ne pas accorder à un personnage du film une position faïtière dans la structure de l'énonciation filmique, on pourrait dire que le film construit ou, au contraire, laisse vide le monde dans lequel se trouve le narrateur. C'est pourquoi il me semble nécessaire, ainsi que je l'ai déjà suggéré à quelques reprises, de revenir sur les catégories de la «voix» selon Genette, d'une part en prenant en compte une dimension propre au cinéma (la possibilité d'établir un acte de référence par visualisation), d'autre part en comprenant le terme «diégèse filmique» au sens de «monde du film».

Il faut alors prendre la mesure des conséquences que notre acception spécifique de la notion de «diégèse» entraîne quant à la signification des substantifs qui en sont dérivés: en tant que «monde», elle est totalement distincte de l'«histoire». Par conséquent, l'opposition extradiégétique *vs.* intradiégétique ne peut plus être un «fait de niveau» (narratif) ainsi que l'envisage Gérard Genette<sup>659</sup>, mais renvoie, comme l'indiquent les préfixes, aux notions d'inclusion/exclusion d'une entité par rapport à un monde donné. Il s'agit d'ailleurs là d'un sens que l'on adopte fréquemment dans les analyses de films. Quant au second couple de notions (homodiégétique *vs.* hétérodiégétique), il permet de préciser l'accessibilité ou l'inaccessibilité qui détermine la relation entre deux mondes. L'apparition du narrateur en tant que personnage dans son propre «récit» indique que le monde de ce récit lui est accessible. Si l'on prend en compte le critère de la visualisation et que l'on croise les deux couples oppositionnels, on constate que seul un narrateur hétérodiégétique peut être considéré comme extradiégétique. En effet, afin de ne pas assimiler tout locuteur *over* à une instance extradiégétique du fait de sa nécessaire non-visualisation, il faut s'interroger sur la relation présumée que ce locuteur entretient avec les «corps» visibles à l'écran. Il résulte de cette comparaison qu'un narrateur homodiégétique est présent à l'image (à de très rares paralipses près): s'il n'est pas visualisé comme narrateur, il l'est néanmoins comme personnage, ce qui lui confère indirectement une certaine consistance qui, nécessairement, l'intègre dans une diégèse. D'ailleurs, sa voix est généralement affectée d'un «grain» plus marqué, car il lui incombe de connoter certains traits de la psychologie d'un personnage. La seule relation possible d'un

659. G. Genette, *Nouveau discours du récit*, *op. cit.*, p. 55.

narrateur homodiégétique avec la diégèse relève donc de l'inclusion (intradiegéticité). Par contre, lorsqu'il s'agit d'un narrateur *over* hétérodiégétique qui n'est pas présenté *in* dans une situation-cadre, celui-ci peut être considéré comme extradiégétique.

Il ne s'agit pourtant pas de remplacer le premier couple de notions par le second, car un narrateur hétérodiégétique peut également être intradiégétique, soit parce qu'il est visualisé en tant que narrateur ou en tant que personnage dans une situation-cadre (cette intradiégéticité peut également être appréhendée rétrospectivement), soit parce qu'il s'agit d'un narrateur qui raconte *in* un récit dans lequel il n'intervient pas. On comprend dès lors les raisons de «la confusion qui s'établit fréquemment entre la qualité d'extradiégétique [...] et celle d'hétérodiégétique» relevée par Genette<sup>660</sup> : en fait, le problème du partage entre ces deux catégories résulte plus fondamentalement du fait que le narratologue n'opère aucune distinction entre diégèse et récit. Ainsi, parmi les diverses anachronies qu'il identifie, les seules qui puissent être le fait d'un narrateur extradiégétique sont les prolepses et les analepses *hétérodiégétiques*, c'est-à-dire celles «portant sur une ligne d'histoire, et donc un contenu diégétique différent de celui (ou de ceux) du récit premier»<sup>661</sup>. Car, pour autant que les événements contés appartiennent au «récit premier», le monde possible qui s'établit est identique à celui qu'occupe le narrateur. De ce fait, ce dernier ne peut être qu'intradiégétique. Appliquée au cinéma, la notion de «*récit premier*» révèle en outre une nouvelle confusion entre diégèse et récit, mais dans le sens inverse de ce que l'on trouve d'habitude chez Genette. En effet, le film peut se contenter de nous montrer un narrateur en situation d'énonciation sans pour autant comprendre cette activité dans une structure narrative (dans le cas de ce que j'ai proposé d'appeler une «situation-cadre»); cette visualisation contribue pourtant bien à poser la (ou une) diégèse.

L'intradiégéticité du narrateur connaît néanmoins des exceptions lorsque la distance temporelle induit des changements interprétables en termes de lois diégétiques. Dans ce cas, c'est donc la *portée* de l'anachronie qui importe, car elle peut occasionner des fluctuations relatives aux états de choses. Si un narrateur relate ce

660. *Idem*.

661. G. Genette, *Figures III*, *op. cit.*, p. 91.

qu'il a vécu deux jours auparavant, l'état du monde et ses propres attributs n'auront probablement pas subi de grands changements; si, par contre, il évoque sa jeunesse, les attributs qui lui sont conférés en tant que personnage du récit enchâssé ne seront pas identiques à ceux de la situation-cadre (par exemple, sa voix *in* pourra être différente de la voix-over). Par la répétition de motifs qu'elle met en place, la trilogie des *Retour vers le futur* (*Back to the Future*, Robert Zemeckis, 1985, 1989 et 1990) exhibe cette interaction entre temporalité et logique modale. Prenons l'exemple de la planche à roulettes du héros, attribut typique du public visé par le film dans le « monde actuel » (l'adolescent américain des années 80). Cet objet fait office d'invariant sur le plan du signifié: dans le premier épisode où le jeune homme voyage dans le temps jusqu'aux années 50, il construit à la surprise de tous un skate à partir d'une trottinette, introduisant un élément inexistant à cette époque; dans le second, il découvre dans le futur un skate volant qu'il emporte avec lui dans le monde des *fifties*; dans le troisième volet de la série, il utilise le même objet au temps de la conquête de l'Ouest. On voit bien qu'il est possible de considérer un degré d'écart entre les mondes, quand bien même ceux-ci ne sont séparés que par un décalage temporel.

Cette corrélation entre temporalité et fictionalité rejoint alors, d'un point de vue logique, l'association que nous avons relevée entre distance temporelle et fictionalité: le rejet dans le passé s'accompagne souvent d'une distorsion des qualités de l'univers considéré, comme le montrent les films qui thématisent le souvenir ou l'oubli. Dans son introduction à l'édition de *L'Année dernière à Marienbad*, Alain Robbe-Grillet déclarait: «Ayant admis le souvenir, on admet sans peine l'imaginaire.»<sup>662</sup> Le film réalisé par Resnais le prouve. D'ailleurs, le cinéaste et le scénariste n'envisageaient pas de la même manière le degré de fictionalité des événements censés s'être produits «l'année dernière», ainsi qu'ils le montrèrent dans les conférences de presse qui suivirent la sortie du film<sup>663</sup>. Il s'agit certes là d'une «fiction dans la fiction» qui multiplie les pistes interprétatives, mais le principe est identique à celui que suppose toute création d'un monde.

662. Alain Robbe-Grillet, *L'Année dernière à Marienbad*, Paris: Éditions «J'ai lu», 1988 [1<sup>re</sup> édition 1961], p. 14.

663. Cité par S. Liandrat-Guigues et J.-L. Leutrat, *op. cit.*, p. 84.

La distance entre les mondes s'accroît dans le cas de prolepses effectuées à partir d'un « point zéro » situé à une date contemporaine de la réalisation du film. Dans ce cas, le futur rejoint l'hypothétique de la science-fiction qui, nécessairement, implique la transformation, généralement régie par des lois intertextuelles, d'un grand nombre d'états de choses. Dans *Terminator* (James Cameron, 1984), le récit des exploits de John Connor conté par l'homme envoyé du futur auprès de Sarah (interprétée par Linda Hamilton) embraye grâce à une voix-explication synchrone puis, lorsqu'il prend fin, laisse entendre rétrospectivement que l'interlocutrice a rêvé ce qui a été audiovisualisé. Cet enchaînement implique l'interaction de deux mondes qui relèvent de sous-genres filmiques différents (le film d'action en milieu urbain contemporain et le film de guerre de science-fiction), notamment en ce qui concerne la distance qui les sépare du monde « réel ». Lorsque, dans le prologue de *Terminator II: Le Jugement dernier* (James Cameron, 1991), une voix-over que nous identifions comme étant celle de Linda Hamilton se fait entendre sur les images d'une terre désolée du futur, nous savons par inférences intertextuelles que ce personnage est « extradiégétique » par rapport aux images : elle appartient à une autre « partie » de la diégèse filmique (qui, en l'occurrence, ne s'arrête pas à un seul film, mais s'étend à ce jour sur trois épisodes). On peut tirer deux conclusions des remarques qui précèdent : d'une part, on qualifiera un narrateur d'« extradiégétique » s'il ne connaît aucune visualisation, ni en tant que conteur, ni en tant que personnage agissant ; d'autre part, la diégèse filmique est susceptible de contenir plusieurs mondes. Ce second point, qui nécessite de considérer d'éventuels cas de *pluridiégéticité*, se répercute sur la façon dont on conçoit le caractère « extradiégétique », puisqu'il s'agit dès lors de savoir par rapport à quelle diégèse ce statut doit être déterminé.

On peut illustrer cette nécessité de redéfinir les catégories genetiennes avec *Avalon* (Mamuro Oshii, 2001), un film qui, en construisant différents mondes, soulève le problème du caractère « extradiégétique » de la musique. Le long métrage d'Oshii « raconte » comment une jeune femme nommée Ash, héroïne d'un *wargame* illégal de réalité virtuelle qui se joue en réseau, tente d'accéder à des niveaux supérieurs puis secrets du jeu dans lequel elle pénètre. Apparaissant dès le générique, la musique d'*Avalon* est indubitablement *over* puisque rien ne laisse penser qu'elle est



entendue par les personnages. Les paroles chantées par le chœur ont des implications narratives indéniables (elles comprennent d'ailleurs le titre du film), fournissant au spectateur certaines clés de ce récit ésotérique bien avant que les personnages ne les possèdent (comme une voix-over explicative de jeu vidéo, mais sous la forme hautement « artistique » et « anti-machinique » de l'orchestre philharmonique, avec chœur et cantatrice soliste). Planant sur des séquences itératives relatant le quotidien sordide de la joueuse, ces chants qui évoquent la terre promise d'Avalon semblent appeler le personnage vers un au-delà que le spectateur est dans un premier temps incité à considérer comme une référence mythologique décontextualisée, un espace extérieur à la diégèse. Nous comprenons toutefois par la suite que cette musique n'est pas « extradiégétique » au sens de Genette, mais provient d'un autre monde, inclus dans la diégèse du film selon la logique d'une hiérarchisation en « classes ». En effet, après avoir franchi une nouvelle étape, Ash passe d'un monde qui nous était présenté dans une image noir/blanc, monochrome (sépia pour les séquences dans le jeu) ou ne faisant qu'un usage très sélectif de la couleur, à un monde en couleurs correspondant au « monde actuel » (ou « Monde d'Origine », MO) du spectateur : elle est censée assassiner un « Non-Revenu » dans une salle de concert où, justement, l'orchestre joue *in* la musique que nous avons entendue *over* pendant plus d'une heure de film. Cette musique est donc intradiégétique par rapport à la diégèse filmique – que je propose de nommer dans de tels cas *mégadiégèse* – et extradiégétique par rapport à la première diégèse ou M1 (Monde 1). Les voix du chœur assurant l'accessibilité d'un monde à l'autre, elles méritent d'être qualifiées, tout comme certains personnages, d'homodiégétiques. Le statut de la voix par rapport aux images dépend de la configuration interne de la mégadiégèse. Dans *Avalon*, M1 se scinde pour donner lieu au M2 du *wargame*, dont l'accessibilité est motivée soit technologiquement par l'existence de « portails » informatiques, soit par l'attitude prépositionnelle du personnage principal (Ash se souvient d'événements qui se sont produits dans le jeu et qui nous apparaissent en *flash-backs*). Dans M1 et M2, la voix chantée se manifeste *over* ; dans M3, elle est jouée *in* (ce que soulignent des gros plans sur la cantatrice) ou *off* (l'enjeu narratif ne se résolvant pas à l'intérieur de la salle de concert, mais dans un espace contigu). De façon plus générale, le film procède à une différenciation et à une interpénétration

des mondes qui s'appuient sur d'autres éléments que ces vocalises éthérées ou que l'usage singulier de la couleur: le monde extrêmement peu peuplé (rues et tramway quasi vides) et meublé de façon rudimentaire (dépouillement extrême des lieux, les systèmes informatiques étant eux-mêmes présentés comme *low tech*) est remplacé dans le monde en couleurs par des plans tournés dans une grande ville contemporaine. Par rapport à MO et à M3 – le film ne nous donne pas la possibilité de les distinguer, de sorte que le spectateur appliquera le principe de l'écart minimal, la représentation d'autres mondes au degré supérieur de fictionalité affiche donc l'«incomplétude» que l'on attribue généralement à ce type d'univers<sup>664</sup>. Oshii, cinéaste issu du milieu du dessin animé, construit entièrement son film comme s'il partait d'une page blanche, utilisant l'image de synthèse pour opérer des soustractions.

En exemplifiant succinctement la théorie des mondes possibles avec *Avalon*, un film qui met en scène une interaction entre deux mondes, à l'instar de *Dark City* (Alex Proyas, 1998) ou de *Matrix* (Andy et Larry Wachowski, 1999) – le récit de ce dernier s'inspirant explicitement du célèbre film d'animation du même Oshii, *Ghost in the Shell* (1995), je choisis bien sûr ce qu'il convient d'appeler un «bon objet». On peut toutefois se demander dans quelle mesure ce ne sont pas au contraire les objets d'étude qui construisent la théorie. Il a fallu la filmologie – et donc le cinéma – pour repenser la notion de diégèse; aujourd'hui, les productions audiovisuelles comme les films qui recourent à l'image de synthèse ou les jeux vidéo me paraissent partiellement sous-tendre l'intérêt suscité par les «mondes possibles» dans les études contemporaines sur la fiction. Dans le champ du cinéma, on pourrait dire que l'«ontologie de l'image photographique» selon Bazin a été supplantée par la question de l'ontologie des univers fictionnels, dont l'étude conduit à détrôner la préséance du «réel» pour se recentrer sur les relations intermondaines. Dans un jeu vidéo – et surtout dans ceux qui reposent sur le principe d'une exploration d'un univers non systématiquement soumise à l'exécution d'une mission (autrement dit à une narrativisation) – le joueur se voit projeté successivement dans une série d'*environnements* (de diégèses incluses dans la «mégadiégèse» désignée par le titre du jeu) dont il doit faire l'apprentissage des lois (qui, littéralement, sont gérées par un

664. Voir Th. Pavel, *op. cit.*, p. 135.

«programme»), notamment quelquefois par le biais de voix-over. De façon bien plus profonde que le seul *zapping* télévisuel permettant de brusques changements de mondes (en quelque sorte intégré dans les films *Une nuit en enfer* ou *Kill Bill*, et dont Cronenberg imaginait les prolongements «modaux» dans *Videodrome* en 1982), les représentations virtuelles ont aiguisé la sensibilité du spectateur contemporain de cinéma envers une appréhension en termes de «mondes possibles». Certes, les surimpressions, maquettes et rétroprojections de *King Kong* (Merian Cooper et Ernest Schoedsack, 1933) permettaient déjà de greffer l'inexistant dans un environnement pseudo-factuel, et les animaux préhistoriques n'ont pas attendu Spielberg pour apparaître sur les écrans. Toutefois, ce qui a changé, ce sont les liens que de tels films peuvent instaurer avec le «monde d'expérience» du spectateur, surtout lorsqu'ils sont visionnés sur un écran qui accueille par ailleurs des jeux vidéo ou permet d'accéder aux «mondes» d'internet. Les échanges intertextuels qui s'opèrent entre un film comme *Star Wars: A New Hope* (George Lucas, 1977) et le jeu vidéo *Rogue Leader* sur Playstation 2 ne font qu'explicitier ces connexions<sup>665</sup>.

Les premières avancées de cette théorie de la fiction dans le domaine francophone sont d'ailleurs à peu près contemporaines de l'un des premiers films qui «réalisent» à proprement parler la jonction entre l'univers d'un film et celui d'un jeu vidéo: *Tron* de Steven Lisberger (1982), production Disney qui raconte ce qu'il advient d'un concepteur de programmes informatiques happé dans le monde d'un jeu électronique. Le rapprochement que je propose ici entre l'avènement d'une théorie et un certain contexte culturel dans lequel le cinéma tient une place prépondérante n'est peut-être pas qu'une «fiction» d'historien du cinéma. Preuve en est la référence que l'on trouve chez l'un des principaux théoriciens de ces questions: Umberto Eco motive en effet la pertinence de l'approche qu'il prône en citant le film *Qui veut la peau de Roger Rabbit?* (*Who Framed Roger Rabbit?*, Robert Zemeckis, 1988)<sup>666</sup>, dont l'intrigue située dans le milieu du cinéma donne lieu à une interpénétration constante de deux mondes à partir du «contrat» posé initialement avec le spectateur postulant que les *toons* ont, en dehors des dessins animés, une existence effective. Or il y a eu

665. Voir Laurent Jullier, *Star Wars. Anatomie d'une saga*, Paris: Armand Colin, 2005, pp. 183-186.

666. U. Eco, *Les Limites de l'interprétation*, op. cit., p. 217.

pléthore de ce type de films durant ces quinze dernières années (récemment encore, une situation identique à celle du film de Zemeckis est exploitée dans une réalisation de son confrère Joe Dante: *Looney Tunes: Back in Action*, 2003), et cela même en dehors des superproductions asiatiques ou hollywoodiennes (par exemple dans *I Want to Go Home* d'Alain Resnais, 1989, où les personnages dessinés font partie du « monde intérieur » des protagonistes humains). La référence d'Umberto Eco me paraît révélatrice de la constitution du cinéma en « bon objet » de ce type de théorie. Or, comme les jeux vidéo ou internet, ce médium fait intervenir la voix, dont l'émanation pose toujours la question de son origine, et donc de son ancrage dans un monde. Étant associée à l'humain, la voix introduit alors un mode d'accès à des univers générés informatiquement qui conditionne fortement le rapport de l'utilisateur à l'instrument technologique.

En fait, la plupart des jeux vidéo réactivent le paradigme de la « machine parlante » qui s'est constitué sur des bases scientifiques au début de l'ère industrielle et qui a culminé dans l'hégémonie de la voix synchrone du cinéma parlant. Il importe par conséquent d'accorder une place particulière à ce facteur anthropomorphisant qu'est la voix. Si la musique d'*Avalon* peut être qualifiée d'« homodiégétique » au même titre que la voix d'un narrateur fictionnel, il n'en reste pas moins que la voix-over possède une double spécificité: sa « matérialité vocale » qui fait d'elle le prolongement désincarné d'un *individu* (en dépit de toutes les *divisions* que dessinent les « frontières » entre les mondes du film) et l'usage du langage verbal, qui déclenche des processus référentiels. Lorsqu'on parle d'intradiégéticité ou d'homodiégéticité, on se réfère à la place qu'occupe la source présumée de la voix dans le monde (ou dans l'un des mondes) du film.

L'examen des rapports entre diégèse et voix-over nous a conduit aux deux composantes du pôle humain que nous n'avons cessé de dialectiser: le corps et la voix. En fait, la voix-over permet la manifestation de la *personne* (grammaticale et humaine) dans un fonctionnement sémiotique fondamentalement impersonnel, mais qui s'humanise à travers son appropriation par le spectateur, et cela d'autant plus fortement que ce dernier associe la production signifiante à un message que lui adresse « quelqu'un ». Car la « coopération textuelle » d'Umberto Eco prend un tour particulier dans le cas de la voix-over: cette dernière verbalise partiellement le « discours intérieur » du spec-

tateur tout en créant un monde qui lui est destiné. Le sentiment d'être en présence d'une instance vocale (donc humaine) perçue comme l'instigatrice de la diégèse filmique dépend de plusieurs facteurs: la détermination d'une relation de subordination verbo-ictonique en faveur de l'instance verbale; l'humanisation que peut permettre la visualisation d'un narrateur homodiégétique ou, plus simplement, le grain de la voix; la place syntagmatique des manifestations *over* (privilège des situations inaugurales); l'importance jouée par la voix dans l'accessibilité du spectateur à la diégèse (ou à l'une des diégèses), etc. En fait, un examen plus poussé de ces questions exige d'une part de traiter la question des actes de référence, d'autre part d'entrer dans une problématique énonciative.

Le texte proféré contribue, lui aussi, à l'établissement d'un monde. C'est pourquoi il est nécessaire de considérer l'interaction entre les référents verbaux et les référents iconiques. Trois cas de figure peuvent se présenter. Primo, la *coréférence verbo-iconique*: le spectateur établit la convergence des deux actes de référence au sein d'un même «plan de référence» qui constitue un monde possible. Lorsqu'elle est différée, cette coréférence souligne le primat de la voix tout en exhibant le principe technique de la synchronisation de mise en chaîne. Simultanée, elle donne l'impression d'une fusion entre le vu et le dit. Secundo, l'*intersection référentielle*, c'est-à-dire la compatibilité du référent verbal avec le «plan de référence» iconique (mais non avec un référent isolé de l'image), souvent établie par le spectateur sur la base d'une déduction logique. Il y a en effet nombre d'informations qui, transmises uniquement par la voix-over, ne sont jamais visualisées, mais néanmoins intégrées par le spectateur à la même diégèse que celle à laquelle appartient l'image. Au début de *Guernica* (Alain Resnais, 1950) par exemple, il n'est pas besoin que les œuvres de Picasso filmées par le cinéaste figurent ce qui est évoqué par la voix-over masculine («les avions allemands au service de Franco bombardèrent Guernica») pour que le spectateur intègre à la diégèse cette information nodale. D'ailleurs, deux éléments de la piste-son (des vrombissements de moteurs et des explosions) créent une situation de coréférence avec le verbal (/avion/ et /bombarder/). De telles références verbales autonomes conduisent bien sûr à placer l'énonciateur verbal fictif dans une position de supériorité par rapport au reste du «matériau» de la représentation. Tertio, la *disjonction référentielle*: ici, nous

rencontrons un cas de pluridiégéticité qui dissocie la diégèse bâtie à partir de référents verbaux autonomes de celle qui est visualisée. On trouve ce cas de figure dans *Son nom de Venise dans Calcutta désert* (Marguerite Duras, 1976) ou dans *Ouvriers, paysans (Operai, contadini, Straub et Huillet, 2001)*. Dire du film de Duras que la voix y est «extradiégétique» au sens de Genette serait méconnaître le potentiel de création de monde que possède le langage verbal: il y a là plutôt deux diégèses distinctes au sein d'une mégadiégèse qui s'établit dans l'imagination du spectateur. Ces constructions diégétiques radicales révèlent l'inadéquation du modèle littéraire à un médium qui conjugue le texte et l'image. Ne jetons cependant pas le bébé avec l'eau du bain: la reconfiguration que j'ai proposée ici confirme à mon sens l'intérêt des catégories de Genette. Par contre, elle démontre que l'application de ces dernières au cinéma doit éviter tout plaquage servile.

Les matières de l'expression filmique se répartissent en deux grandes catégories: les signes iconiques et les manifestations verbales. En dépit de l'importance du rôle endossé par le verbal depuis les projections cinématographiques des premiers temps, et de certaines tentatives «expérimentales» qui ont consisté à supprimer l'image ou à minorer son rôle tout en s'inscrivant encore dans l'institution du cinéma (Ruttman, Lemaître, Duras, etc.), l'image demeure première dans la conception que le public se fait traditionnellement du spectacle cinématographique. La dimension sonore, notamment le verbal oral, est considérée comme facultative, surtout lorsqu'elle est *over*. Ce postulat, discutable lorsqu'il conduit à négliger la place du son dans la représentation audiovisuelle, doit être pris en compte dans l'analyse des interactions verbo-iconiques dans la mesure où il nous conduit à privilégier l'apport fourni par le verbe dans la réception de l'image, plutôt que l'inverse. Il ne s'agit bien sûr pas d'établir une relation de subordination: on sait que le flux d'images peut tout aussi bien se calquer sur une piste-son préexistante (par exemple dans les *clips* vidéo, ou dans des séquences de longs métrages construites sur ce principe), et que cet apport verbal ne s'évalue pas en termes de valeur ajoutée (la somme des informations iconiques et verbales), mais transforme profondément le contenu sémantique de l'ensemble du message (le résultat de la conjonction des deux types de signes est, pour reprendre la comparaison mathématique, le *produit* de l'un et de l'autre). L'expérience de Chris Marker consistant à appliquer

successivement trois énoncés verbaux fort différents du point de vue sémantique sur une même série de plans dans *Lettre de Sibérie* (1958) a montré l'ampleur possible de telles transformations, même s'il ne faut pas négliger, sous peine de tomber dans le poncif de l'implacable polysémie de l'image que seul le verbal parviendrait à canaliser, la capacité que possède l'image à produire par elle-même une signification passablement univoque<sup>667</sup>.

Le caractère facultatif de certaines configurations vocales et verbales m'incite à concentrer mes remarques sur la façon dont le verbal (oral et *over*) influence la perception de l'image, et non l'inverse. En outre, j'écarterai l'image non figurative qui implique une divergence trop importante avec le fonctionnement référentiel du langage verbal: alors que celui-ci ne peut éviter de référer, l'image cesse de le faire lorsqu'elle se situe en dessous d'un degré minimal d'iconicité.

On considérera deux signes comme «coréférents» lorsque le référent de chacun d'eux, c'est-à-dire l'objet du monde (qu'il existe ou non, qu'il soit concret ou abstrait) extérieur à ces signes, est identique. La coréférence est bien sûr à l'œuvre à l'intérieur des deux supports sémiotiques considérés (image photographique animée et langage verbal): on peut l'utiliser pour parler d'un élément visuel qui réapparaît d'un plan à l'autre (par exemple le visage d'un personnage, qui est souvent le garant de la cohérence diégético-narrative) ou, comme le font les linguistes, pour désigner le statut d'un pronom qui se réfère au même objet qu'un substantif compris dans le même énoncé. L'emploi que je propose de la notion de «coréférence» ne correspond pas à cet usage courant puisque je la concevrai toujours au sens de «coréférence verbo-iconique», afin de rendre compte, sur un plan sémantique, de la relation *verticale* qui s'instaure entre les composantes visuelles et sonores internes à la représentation filmique: un syntagme nominal du texte se réfère à un objet du monde visualisé à l'écran.

On distinguera en outre la *référence* de la *signification*: dans la conjonction du mot et de l'image, il y a en effet toujours quelque chose qui excède le référent ou en médiatise la reconnaissance, puisque les signifiés des signes iconique et linguistique sont multiples. Même dans le cas le plus élémentaire d'une convergence entre le référent d'un mot et celui d'une représentation visuelle, ce surplus existe (ne serait-ce qu'en vertu du degré d'abstraction du signifié),

667. Martine Joly, *L'Image et les signes*, Paris: Nathan, 1994, pp. 81-88.

et se voit décuplé lorsque interviennent des phénomènes de connotation. C'est pourquoi, nécessairement, l'image et le mot s'enrichissent mutuellement. En ce qui concerne le seul langage verbal, on constate par exemple que les synonymes (qui possèdent un référent partiellement identique) ne possèdent jamais exactement la même valeur (dénotative et connotative). Ainsi, le choix d'un terme plutôt que de l'un de ses synonymes influence inévitablement le sens global que l'audiospectateur confère à la représentation. C'est pourquoi la notion de « redondance » ne doit pas être conçue de façon péjorative, ainsi que le fait communément la critique cinématographique afin de motiver le sentiment d'« inutilité » éprouvé à l'égard de certaines voix-over. La redondance ne doit pas être réduite à l'emploi pléonastique banni des manuels scolaires d'apprentissage de la dissertation, mais appartient de plein droit, comme l'explique Henri Suhamy, à la catégorie des figures de style :

« La redondance garde la même idée sous plusieurs formes [...]. On ne peut pas dire qu'il y ait dans aucune langue deux mots ou locutions complètement identiques et superposables, de même dénotation et connotation. La redondance ajoute toujours une nuance, et le rythme, le mouvement obtenus par ce moyen appartiennent au domaine des figures. »<sup>668</sup>

On comprend que le constat effectué par Suhamy à l'intérieur d'une même langue est d'autant plus patent lorsqu'on a, comme au cinéma, juxtaposition de deux supports sémiotiques. Suhamy pointe par ailleurs la dimension rythmique du phénomène qui concerne également la temporalisation induite au cinéma par la synchronisation.

L'étude de la coréférence permet de comprendre comment la « voix de la machine » s'avère capable de créer un monde, à l'instar d'un conteur qui nous livrerait un récit oral, mais en proposant (et en imposant) un réglage strict des processus sémantiques effectués au niveau du déroulement temporel par la synchronisation. En effet, la dimension référentielle du discours audiovisuel fait office d'instrument intersémiotique de *construction d'un monde* qui, à ce niveau, n'exige pas d'intervention humaine, mais résulte d'un fonctionnement textuel et de l'activité cognitive de l'audiospectateur. Ces processus sémantiques dépendent bien sûr fondamentalement des contraintes de la synchronisation.



La situation basique de la *coréférence verbo-iconique simultanée* est la suivante: par le biais de syntagmes nominaux, le texte *over* aligne, dans un temps donné, une série de référents; dans le même laps de temps, l'image livre de son côté un nombre très important de signes visuels dont la plupart possèdent un référent. Si l'on compare les données issues de ces deux sources, on observe inévitablement, de part et d'autre, un débordement: parmi les référents iconiques, seuls quelques-uns peuvent être considérés comme identiques à certains référents verbaux proférés au moment de leur apparition à l'écran. Inversement, il y a des unités verbales dont le référent n'est pas visualisé. Ce sont néanmoins les processus coréférentiels qui assurent le lien sémantique entre l'image et la voix-over, comme le font à un niveau perceptif les points de synchronisation.

Il faut noter que les rapports entre l'image filmique et la parole ont souvent été traités du point de vue des capacités métalinguistiques du verbal. Cette tradition remonte au premier écrit consacré par Christian Metz en 1965 à la sémiotique, dans lequel il rabattait la distinction parole-ancrage *vs.* parole-relais de Barthes sur l'opposition de ce qu'il considérait comme les deux grandes formes de manifestation du langage verbal au cinéma, la parole diégétique et la parole «non-diégétique». La définition de Metz est révélatrice des premières conceptualisations suscitées par le procédé de la voix-over antérieurement à l'hégémonie de la narratologie genettienne, qui contribua à éloigner les théoriciens du cinéma de la comparaison intersémiotique proprement dite: «La parole intervient en fonction de métalangage par rapport aux signifiants non verbaux; elle appartient encore à la lexie [à la même «unité de lecture»], mais elle la «coiffe» plutôt qu'elle n'en fait partie.»<sup>669</sup> Si la parole *over* «coiffe» l'image, c'est dans la mesure où elle propose également un référent qui se conjugue à celui de l'image, le complète ou s'y oppose. Il importe par conséquent d'examiner en priorité le fonctionnement de la coréférence verbo-iconique.

Si la voix-over apparaît comme un objet pertinent pour analyser l'influence du verbal sur la lecture iconique à travers les phénomènes coréférentiels, c'est parce que le rapport qu'elle postule

668. Henri Suhamy, *Les Figures de style*, Paris: PUF, 1981, pp. 67-68.

669. Christian Metz, «Les sémiotiques (À propos des travaux de Louis Hjelmslev et d'André Martinet)», *Communications*, N° 7, 1966, repris in *Essais sémiotiques, op. cit.*, p. 12.

entre le mot et l'image permet à ces phénomènes sémiotiques d'y être particulièrement manifestes. Plus la voix est ancrée dans l'image, plus la diversité des phénomènes de coréférence simultanée diminue. C'est pourquoi la voix synchrone correspond au spectre le plus réduit des possibilités, surtout lorsque le film nous montre le locuteur en gros plan. Dans un tel cas, aucune coréférence ne s'opère entre l'image et les paroles prononcées par le personnage de l'écran, à l'exception de l'embrayeur « je » ou de l'auto-désignation (lorsque le locuteur évoque son propre visage). Le lien entre ce que l'on voit et ce qui est dit est par contre de l'ordre de la signification : une relation peut s'établir entre le contenu des paroles et la mimique du locuteur, interprétée par le spectateur selon des codes extracinématographiques établis socialement. S'il n'y a pas de liaison, l'expression faciale devient alors un signifié de connotation du verbal. Une situation de coréférence peut néanmoins s'instaurer si le locuteur imite physiquement ce qu'il évoque verbalement : par exemple, il peut se dire « dégoûté » tout en exprimant le dégoût sur son visage, ou imiter par sa gestuelle une personne, une action, la forme d'un objet, etc. dont il fait mention dans son discours. On comprendra néanmoins que le potentiel de coréférentialité est largement inférieur à ce que l'on trouve dans les séquences en voix-over : moins l'image est asservie à la monstration du locuteur, plus les interactions entre le référent de ses paroles et l'image sont susceptibles d'être nombreuses. Le désancrage de la voix, lorsqu'il ne s'effectue pas sous la forme de la déliaison vocale, permet une meilleure exploitation de la coréférence simultanée, sans compter que le locuteur synchrone, en tant que partie intégrante de l'univers filmique, dispose de moins de possibilités référentielles en raison des exigences de la cohérence diégétique.

Il faut préciser que la coréférence n'implique pas nécessairement une identité de tous les référents visuels et verbaux. D'une part, certaines références verbales, que je propose de nommer *autonomes*, ne connaissent aucun « équivalent » visuel, mais s'inscrivent néanmoins dans la cohérence diégétique (une fois asserté verbalement, ce référent « existe » dans le monde du film). D'autre part, la relation qu'entretient le mot avec l'image peut aussi reposer sur une absence de correspondance qui connaît des degrés variables. Dans son analyse de la coréférence verbale, Jean-Claude Milner relève deux types de « distinction référentielle » : la disjonction (aucun trait commun entre deux références) et l'intersection

(parenté partielle)<sup>670</sup>. Au sens strict, l'*intersection référentielle* ne paraît pas applicable aux relations verbo-iconiques. Milner prend en effet l'exemple des deux termes «jeune fille» et «secrétaire». Si l'on imagine un tel cas pour les relations entre image et verbal, il n'est pas possible d'opérer une distinction entre l'intersection référentielle et la coréférence, car le référent visuel n'est pas suffisamment particularisant. On pourrait toutefois entendre cette notion dans un sens large, moins spécifiquement lié au découpage lexical du langage. L'«intersection» sert alors à qualifier les inférences diégétiques ou encyclopédiques<sup>671</sup> qui permettent au spectateur de déduire la présence du référent concerné dans l'espace visible sur la base d'un élément appartenant au même champ sémantique. Le principe de cette déduction est d'ordre métonymique, et permet de suspendre les différences existant entre les deux supports sémiotiques en termes de particularisation et de catégorisation. Si, par exemple, la voix-over mentionne «la police» alors que nous voyons un immeuble qui présente tous les traits d'un «commissariat», ou «la nourriture» alors que l'image représente une pomme, il n'y a pas à proprement parler de coréférence mais, comme le dit Milner à propos de l'intersection, l'un des termes peut «satisfaire aux exigences» de l'autre. Quant à la *disjonction référentielle*, elle s'avère à mon sens utile pour rendre compte de deux fonctionnements possibles de la représentation audiovisuelle: ou bien l'image, non figurative, ne permet pas l'identification d'un référent, ou bien son référent est en tous points (ou en de nombreux points) incompatible, à l'échelle globale du «plan de référence»<sup>672</sup>, avec le référent verbal. La disjonction référentielle se distingue de la référence autonome par le fait qu'elle postule la présence de deux termes possédant deux référents différents, alors que la référence verbale autonome ne peut être corrélée à aucun référent visuel. Le cas de la disjonction, qui interdit toute mise en relation entre l'image et le verbal, est fort rare, et peut-être même inexistant dans l'absolu si l'on prend en compte les conditions de réception,

670. Jean-Claude Milner, *Ordres et raisons de langue*, Paris: Seuil, 1982, pp. 12-13.

671. À la suite d'Umberto Eco, je qualifierai d'«encyclopédiques» toutes les inférences qui ont trait aux connaissances du spectateur sur le monde (voir U. Eco, *Lector... op. cit.*, p. 33).

672. Pour passer d'une unité discrète (signe iconique, lexème) à un ensemble actualisé de ces mêmes unités (contenu du plan ou d'une série de plans, phrase ou texte), j'utiliserai l'expression «plan de référence», plutôt que celle, proposée par Milner, de «segment de réalité», le terme «réalité» me semblant sujet à confusion dans un médium comme le cinéma, qui dispose généralement d'une forte valeur indiciaire.

puisque l'audiospectateur tentera machinalement de repérer des coréférences verbo-iconiques, fût-ce en forçant le texte ou l'image. Ainsi, on peut dire qu'il suffit qu'une série d'énoncés verbaux corresponde « trop peu » à l'image pour que le spectateur ait l'impression d'une telle disjonction. Il ne s'agit pas seulement d'une caractéristique factuelle du texte filmique, mais d'un effet sur le spectateur. Je recourrai donc à la notion de « disjonction référentielle » pour désigner le fonctionnement des relations entre le verbe et l'image dans des films (ou des segments de films) où la voix-over semble avoir été arbitrairement plaquée sur une suite d'images, comme cela est le cas dans *Son nom de Venise dans Calcutta désert* (Marguerite Duras, 1976). La disjonction n'est dès lors pas sans conséquences sur la diégèse filmique, car celle-ci se voit en quelque sorte dédoublée<sup>673</sup>.

La relation d'opposition que la disjonction référentielle met en jeu ne ressortit pas forcément de manière exclusive à une séparation nette des deux plans de référence. Les non-coïncidences sémantiques entre le dit et le vu dont certains films usent de façon rhétorique doivent être appréhendées par l'audiospectateur pour être efficaces, c'est-à-dire qu'elles exigent la présence d'un point de repère à partir duquel l'écart peut être évalué. Par exemple, si l'audiospectateur de *L'Homme qui ment* (Alain Robbe-Grillet, 1968) comprend que l'affirmation du narrateur *over* (selon laquelle le café du village est vide) s'oppose à l'image qui nous montre de nombreux clients attablés (disjonction entre les termes /aucun/ et /plusieurs/), c'est parce qu'il existe une base référentielle commune (coréférence: /intérieur/ et /café/). Je propose de qualifier *contraste sémantique* cette disjonction référentielle restreinte à certaines unités.

On peut donc dire que pour chaque référent verbal, le référent iconique « correspondant » est soit *présent* (coréférence), *opposé* (contraste sémantique), *inadéquat* (disjonction référentielle) ou *suggéré* (intersection référentielle). En cas de coréférence, deux paramètres sont susceptibles d'accroître, à partir de la situation minimale de la voix synchrone, la latitude des liens référentiels instaurés entre le verbal et l'image: l'élargissement de l'espace du montré (en variant l'échelle de plans ou en effectuant un mouvement de

673. Voir la notion de « montage asynchrone » chez Marie-Claire Ropars-Wuilleumier, *Le Texte divisé*, Paris : PUF, 1981, p. 159.

caméra) et la suppression de la visualisation du locuteur. Ces deux opérations connaissent naturellement différents degrés de réalisation. Par exemple, comparativement aux autres manifestations vocales, la voix synchrone restreint indubitablement les phénomènes de coréférence, d'autant plus qu'elle est fréquemment associée à une individualisation du locuteur dont le visage est régulièrement filmé en gros plan. Or, plus le cadre est large, la profondeur de champ importante et l'image dense en «objets du monde», plus la probabilité est grande d'obtenir des éléments verbaux dont le référent est visualisé simultanément. Dans la scène de la signature du contrat de *Citizen Kane* (Orson Welles, 1941), la présence de l'enfant qui joue dehors à l'arrière-plan permet de visualiser le référent indirect des paroles (intersection référentielle) : le gosse est littéralement réifié en devenant «l'objet» même de la tractation (sans parler de cette autre «chose» qu'il manipule, la luge la plus célèbre de l'histoire du cinéma, qui, elle, demeure innommée jusqu'à l'avant-dernier plan du film). Ainsi que l'illustre le film de Welles avec ses enchâssements narratifs et les «News on the March» en voix-over, la coréférence simultanée induit généralement l'acousmatisation de la voix.

Bien qu'il s'agisse d'un phénomène sémantique, la coréférence n'en est pas moins soumise au réglage temporel de la synchronisation. Les approches sémiologiques qui, dans les années 70 (Barthes, Lindekens, Metz, etc.), extrayaient du flux du déroulement filmique certains liens ponctuels entre unités lexicales et référents visuels ou dissertaient sur le cinéma à partir de l'image fixe, occultaient à mon sens cette dimension dynamique de la lecture du film. D'ailleurs, elle endosse également un rôle décisif dans le cinéma muet non bonimenté, puisque les intertitres intercalés entre deux plans iconiques devaient se contenter de simuler la simultanéité, contraints d'interrompre momentanément la visualisation d'un personnage en train de parler (ce qui pouvait être fort précis lorsque des mots isolés étaient articulés par les acteurs de façon ostentatoire, comme dans *La Passion de Jeanne d'Arc*, film précisément réalisé en la date charnière de 1928). Contrairement à l'intertitre qui introduit une rupture, la voix-over, instrument idéal d'un pont sonore, linéarise l'enchaînement des plans et influence la temporalité filmique.

Si l'on veut envisager globalement les liens entre coréférence et inscription de la parole dans le déroulement filmique, il faut comparer chaque moment d'énonciation d'un référent verbal avec

l'instant précis d'apparition du référent iconique correspondant sur la chaîne filmique. En prenant comme point de repère l'élément verbal, on obtient trois cas de figure: la mention verbale semblera se «réaliser» dans l'image avec un certain retard (*coréférence rétrograde*), au même instant ou presque (*coréférence simultanée*, avec une certaine marge d'approximation liée au temps de la lecture de l'image) ou avant la monstration du référent iconique (*coréférence différée*).

La coréférence différée est fréquente lorsque la voix est synchrone, puisque l'élargissement du cadre à l'environnement du locuteur – élargissement qui est comme on l'a vu favorable aux interactions verbo-iconiques – s'effectue généralement à la fin du dialogue. De telles paroles diégétiques donnent néanmoins rarement lieu à une coréférence *stricto sensu*: on dira plutôt que la *signification* du plan fait écho à celle du verbal. L'audiospectateur ayant encore à l'esprit les dernières phrases des personnages, une relation plus ou moins lâche s'établit entre leur contenu et le référent visuel qui leur succède, surtout lorsque ces images, silencieuses, constituent une sorte de «caisse de résonance» du dialogue. Prenons, parmi les très nombreux films dans lesquels s'observe ce type d'interactions entre parole et image, l'exemple du plan-séquence de *Gangs of New York* (Martin Scorsese, 2002), qui donne lieu à un ample mouvement d'appareil circulaire passant, sans individualiser aucun personnage, par les différentes «étapes» de l'arrivée des émigrants sur le sol américain, étapes concentrées dans le lieu unique d'un débarcadère. Du boniment du recruteur aux cercueils qui arrivent par le même bateau, les différentes phases de la conscription pour la Guerre de Sécession sont illustrées en un seul plan synthétique (à l'instar d'œuvres picturales qui représentent dans l'espace unique de la toile différentes actions pregnantes d'une même histoire, par exemple la vie ou la Passion du Christ). Ce plan sonore mais dépourvu de voix-action fait suite à une longue harangue où Cutting (Daniel Day Lewis) reproche au politicien, Tweed, le bon accueil réservé aux immigrants. À la fin de cette discussion véhémement, Cutting se retourne et s'éloigne, alors que Tweed lui lance en voix *off*: «Vous tournez le dos à l'avenir.» À partir de cet instant, sans coupe, la caméra pivote sur la gauche et entreprend son long déplacement. La synchronisation de mise en cadre met ainsi en exergue le dernier mot du dialogue, «avenir». Ostensiblement «discursif» en raison de son *finale* ironique (les

dépouilles arrivant par le même navire que les futurs conscrits, «bouclant la boucle» du mouvement circulaire), ce plan répond en quelque sorte aux plaintes de Cutting (c'est-à-dire à l'enjeu global de la conversation), mais surtout explique par l'image ce dont cet «avenir» évoqué verbalement sera fait: l'Amérique se construit sur l'exploitation massive d'émigrés utilisés comme chair à canon. L'expression «tourner le dos» annonce également la fin du film: les héros de l'histoire n'appartiennent déjà plus à leur époque – ils disparaissent en fondu alors que le décor de New York évolue – qui connaît alors de profonds bouleversements sociopolitiques.

Un exemple similaire de réorientation «monstrative» du sens transmis verbalement est traité par Jean Mitry – l'un des rares théoriciens «classiques» du cinéma à s'être penché sur les liens entre parole et image d'un point de vue sémantique, rythmique, etc. – dans *Esthétique et psychologie du cinéma*. L'auteur mentionne un dialogue de *Cavalcade* (Frank Lloyd, 1933) qui a lieu sur le pont d'un paquebot entre deux jeunes mariés. Alors que l'épouse exprime son bonheur en commençant notamment une phrase par «s'il m'advenait de mourir demain», la caméra s'éloigne progressivement des protagonistes en travelling arrière, et, comme le dit Mitry, «laisse voir, à l'instant précis où la jeune épouse vient d'achever sa phrase, une bouée accrochée au bastingage et sur laquelle on lit «Titanic»<sup>674</sup>. Il y a bien simultanéité et contraste, mais ce cas relève, comme dans l'exemple de *Gangs of New York*, de la signification et non de la référence, puisque, selon la tradition linguistique, on ne peut étudier l'acte de référence d'une phrase qui n'est pas nominale. L'épouse n'évoque nullement le transatlantique qui est ensuite visualisé par synecdoque (en ce sens, le principe de «l'intersection» est dédoublé); par contre, l'un des mots de son énoncé (la forme infinitive du verbe «mourir») contient un sème également compris, de façon indirecte, dans l'image – ou plutôt dans le verbal présent à l'image sous la forme d'une mention écrite diégétique: à l'oralité succède l'écrit (comme on dit: «C'était écrit»). Pour que le spectateur comprenne ce que signifie «Titanic» dans ce contexte, il faut qu'il fasse intervenir des inférences encyclopédiques. Dans le cas de ce film, le naufrage de 1912 est suffisamment célèbre pour que le référent de «Titanic»

674. J. Mitry, *Esthétique et psychologie du cinéma*, Vol. 2, *op. cit.*, p. 97.

soit identifié et la signification comprise par la majorité des destinataires, et cela même de nos jours (après le succès mondial du film de Cameron, il n'y a probablement plus beaucoup de spectateurs de cinéma sur le globe qui méconnaissent cette catastrophe). Il peut arriver cependant que le savoir extratextuel exigé pour une bonne conduite des opérations de production du sens ne soit pas (ou plus) partagé par la communauté des «lecteurs». Si le langage ne peut éviter de référer, l'acte de référence peut, lui, conduire à un échec pragmatique.

Dans ces deux films, la signification s'élabore en dehors de l'univers diégétique, en quelque sorte par-dessus les personnages. Cutting ne se rend pas compte de la façon dont son pays exploite les émigrants et la jeune mariée ne sait pas ce qui l'attend. Concrètement, les deux personnages «tournent le dos» à ce que le spectateur a le loisir d'observer. La gestion temporelle de la coréférence conditionne donc dans une large mesure la détermination du type de focalisation.

Il faut noter que la synchronisation à l'œuvre dans les processus de coréférence influe sur les mécanismes de la réception spectatorielle. Alors que la coréférence simultanée agit surtout au niveau de la perception visuelle en pointant un élément dans l'image (c'est l'opération d'«ancrage» selon Barthes), les coréférences rétrograde et différée s'émancipent de ce lien étroit à la représentation visuelle et exigent du spectateur qu'il se souvienne d'éléments montrés ou énoncés antérieurement. Elles endossent donc une fonction anaphorique/cataphorique qui renforce la cohésion du discours filmique, mais influencent moins directement la lecture du plan que ne le fait la coréférence simultanée. Celle-ci tend en effet à orienter le parcours de l'œil, l'audiospectateur essayant d'identifier l'élément susceptible de correspondre à l'un des termes de l'énoncé *over*. Si aucun référent iconique équivalent n'est présent, le spectateur stockera l'information verbale; il la convoquera éventuellement plus tard, dans le cas où ce référent apparaîtrait de façon différée. Cette opération s'effectue de manière plus ou moins consciente en fonction de la familiarité du spectateur avec certains codes narratifs.

Les phénomènes de report ou d'anticipation de la coréférence donnent lieu à des configurations qu'il s'agit d'étudier au cas par cas. Ainsi me semble-t-il absurde de proposer des lois générales du type de celles que propose Jean Mitry dans le texte suivant:



«[...] le texte risque de faire double emploi avec ce que les images montrent nécessairement. Afin d'éviter ces effets pléthoriques, il convient de décaler tant soit peu l'image et le son, de détruire leur production simultanée. Un décalage d'une demi-seconde convient généralement mais tout dépend, bien entendu, de la longueur de la scène, de son caractère et de son importance. Cela peut aller d'une demi-seconde à deux secondes. Toutefois, en règle générale, le commentaire doit intervenir *après* l'image. L'inverse est aberrant; à tout le moins anticinématographique.»<sup>675</sup>

Le commentaire extrêmement normatif formulé par Mitry dans cet ouvrage tardif ne prend en compte la question de la temporalisation du défilement parallèle de la bande-image et des paroles que pour prôner une seule des diverses interactions possibles. Comme souvent chez les théoriciens, l'article de Mitry témoigne d'une confusion entre redondance et synchronisation simultanée, et la subordination de l'image au mot est dépréciée en vertu de sa nature prétendument moins cinématographique. Il est évident que ces a priori esthétiques ne correspondent aucunement à la complexité sémiotique des phénomènes concernés. La détermination du caractère «simultané» de la relation verticale entre les deux pistes – abordée par Mitry dans une tentative d'approximation mathématique fort simpliste en regard des implications de la coréférence – demande par exemple à s'interroger sur les effets produits sur le spectateur.

Dans le sillon des recherches menées par les théoriciens de la Gestalt, certaines «lois» – dont on reconnaîtra non seulement les limites intrinsèques, mais aussi l'insuffisance pour traiter toutes les composantes de la référence verbo-iconique – ont été élaborées par les sémiologues ou les publicistes pour tenter de comprendre et de maîtriser l'influence de la composition de l'image sur le parcours de l'œil du lecteur. La voix-over constitue indéniablement un facteur d'orientation perceptive dans la mesure où elle provoque souvent ce que j'appellerais une *mise en relief* de certains éléments de l'image (notion-clé pour une étude de la coprésence de l'image et du mot). La nomination verbale permet en effet d'isoler l'un des référents de l'image, alors qu'une telle opération n'a pas d'équivalent dans le sens inverse, puisque le verbal est incapable

675. Jean Mitry, *La Sémiologie en question*, Paris: Cerf, 1987, p. 175.

d'actualiser simultanément plusieurs référents. Ce phénomène tient donc à la « polyréférentialité » du message visuel : seul l'un des très nombreux éléments de l'image est identifié verbalement. Claudine de France, théoricienne spécialisée dans l'étude et la réalisation de documentaires anthropologiques, a tenté d'évaluer l'effet du commentaire sur l'image en proposant une opposition (souligné *vs.* estompé) qui rend compte du processus de mise en relief d'éléments visuels par le verbal<sup>676</sup>. Notons toutefois qu'en soulignant un élément visuel, la voix-over dénote certes une intention de l'instance de réalisation, mais rien n'interdit que quelque chose d'autre dans le paysage n'attire l'attention du spectateur, quelle que soit la « neutralité » présumée de cet environnement. Par exemple, le spectateur qui choisit de se concentrer sur le récit fera primer les actions, rejetant le reste de l'image dans l'estompé, alors qu'un historien s'attachera peut-être à la précision de la reconstitution. La dimension narrative de la sélection d'éléments dans l'image peut être associée au sens métaphorique que Harald Weinrich a conféré à la notion de « mise en relief » pour rendre compte, du point de vue de la structure du récit, du rôle joué par le passé simple (temps du premier plan, dont les verbes font saillie en indiquant un moment précis) et par l'imparfait (temps de l'arrière-plan) de la langue française<sup>677</sup>. Cette distinction qui me paraît tout à fait opératoire vaut bien sûr également pour le texte *over*, dont les temps verbaux ne connaissent aucun équivalent dans l'image. Ils participent de façon décisive à la mise en relief narrative.

Si l'on s'en tient à la dimension visuelle, ces mécanismes se compliquent déjà passablement au cinéma par rapport à l'image fixe en raison de la longueur définie de chaque plan qui impose sa durée au spectateur : la lecture « tabulaire » identifiée par les analystes de la bande dessinée<sup>678</sup> y est nécessairement inféodée à la lecture linéaire. De plus, la possible multiplication des mouvements (dans l'image et de l'image) et les rapports résultant de la succession des plans décuplent les facteurs qui interviennent au cours de la lecture. Contrairement à l'image publicitaire fixe qui doit « spatialiser » les éléments déterminant le parcours de la lecture, le moment d'apparition

676. Claudine de France, « Image et commentaire : du montré à l'évoqué », *Hors Cadre*, N° 3, 1985, pp. 146-147.

677. H. Weinrich, *op. cit.*, pp. 112-117.

678. Voir Pierre Fresnault-Deruelle, « Du linéaire au tabulaire », *Communications*, N° 24, 1976, pp. 7-23.

de chaque donnée visuelle ou auditive est prédéterminé au cinéma par le principe de la synchronisation. À la suite des travaux du psychologue américain A. L. Yarbus (dont l'ouvrage *Eye Movements and Vision* paraît en 1967), des expériences oculométriques ont permis de reproduire électroniquement les mouvements oculaires grâce à des techniques (rassemblées sous l'appellation « *visual tracking* ») qui ont connu récemment un certain essor dans les firmes de développement de matériel médical. En tant que composante de ces « visuels » contemporains que sont les sites internet, la voix-over est à compter parmi les objets auxquels s'appliquent ces méthodes d'optimisation de la lisibilité.

On peut donc dire que la voix-over simultanée à la visualisation de l'un de ses référents constitue un paramètre important de l'orientation de la perception visuelle, d'autant plus que le fonctionnement de la vue est géré par l'activité cérébrale. Des études récentes dans le domaine des neurosciences montrent en effet que le traitement cognitif de l'information pilote les facultés perceptives. Bryan Kolb et Ian Wishaw soulignent par exemple le fait que la notion même de « son » n'a d'existence que dans la mesure où intervient le travail cognitif du sujet percevant<sup>679</sup>. Le verbal se présentant dans un film parlant sous une forme orale, le cerveau intervient dès la saisie auditive des mots. L'analyse cognitive de l'information sensorielle n'est cependant pas seule en cause, car le langage verbal entraîne, en tant que support de la communication, une capacité particulière de la part de l'auditeur qui n'est pas suscitée par les autres types de son (si ce n'est dans le cas de la musique, qui offre également une série de sons « signifiants ») : dans le cas de la perception auditive d'unités linguistiques, il est étonnant de constater que « la vitesse de l'entrée dépasse beaucoup la capacité du système auditif à transmettre chaque parole comme une unité isolée d'information<sup>680</sup>. La perception auditive (mais aussi visuelle, qui est encore plus lente) ne peut donc être considérée indépendamment de la compréhension du message verbal. En raison de la rapidité de ce décryptage, la compétence linguistique du spectateur exerce une influence décisive sur la réception globale de la représentation audiovisuelle. Étant guidés par le cerveau, les mouvements oculaires sont nécessairement soumis à

679. Bryan Kolb et Ian Q. Wishaw, *Cerveau et Comportement*, Bruxelles: De Boeck Université, 2002, p. 279.

680. *Ibid.*, p. 325.

l'établissement d'une intelligibilité qui procède, à la vision d'un film à voix-over, d'une comparaison entre le contenu verbal instinctivement et promptement établi et les informations transmises visuellement.

Une analyse satisfaisante des données langagières et sensorielles nécessite une durée suffisante de persistance à l'écran du référent visuel. En effet, pour que l'audiospectateur puisse orienter la lecture du plan en fonction des informations verbales, il faut que le film lui laisse le temps de parcourir l'image après le décodage du message linguistique. Lorsqu'il est question de coréférence «simultanée», il est donc clair qu'il s'agit d'une approximation, mot et référent visuel n'apparaissant que rarement au même point de la bande pelliculaire. Généralement, on observe plutôt une avance du verbal sur l'image, d'autant plus que la perception auditive procède de façon séquentielle, alors que la vue opère une analyse synchrone de différents paramètres de l'image (la différence de vitesse existant entre la propagation des ondes sonores et celle des ondes lumineuses me paraissant ici négligeable).

Il faut préciser que les tentatives de conceptualisation des rapports entre l'activité cognitive et la perception visuelle s'appliquent principalement à des manifestations verbales dont on dira, en termes d'analyse des discours, qu'elles relèvent d'une dominante informative (conjointe à la fonction argumentative dans le message publicitaire). Or les rapports verbo-iconiques traditionnels des films parlants engagent une lecture différente, mue quant à elle par un intérêt esthétique, par une attention sélective dirigée en fonction de compétences narratives et par un désir de fiction et d'identification à un sujet parlant. On retiendra toutefois que la coréférence simultanée est liée à un type d'interaction particulier entre image et parole: le mot identifie un élément dans l'image, il le fait passer à l'avant-plan «cognitif».

Les coréférences rétrogrades et différées permettent quant à elles une certaine autonomisation par rapport à la lecture du plan. Le type de synchronisation qui est majoritairement requis par ces processus sémantiques est donc celui de la mise en chaîne. Le théoricien Iouri Lotman avait prêté attention à ce décalage temporel entre l'apparition des référents iconique et verbal en le rattachant à une forme de «contrepoint orchestral» (notion empruntée à Eisenstein, Alexandrov et Poudovkine). Lotman exemplifie significativement ce point avec un film, *Hiroshima mon amour* (Alain Resnais, 1959), dans lequel la voix-over intervient de façon particulièrement com-

plexe, ainsi que nous le verrons au chapitre VII. Lotman constate qu'au moment où la Française interprétée par Emmanuelle Riva évoque le chat qu'elle aurait aperçu dans la cave, l'animal n'apparaît pas dans le *flash-back*, tandis qu'une image ultérieure nous montre les yeux du félin brillant dans l'obscurité<sup>681</sup>. Les modalités de ce type de traitement temporel de la coréférence sont nombreuses; dans un film comme celui de Resnais, elles exigent une participation active du spectateur. Il importe en effet d'examiner également les effets de ce décalage temporel au-delà de sa seule «fonctionnalité» en tant que processus d'intelligibilité du film, notamment au niveau de l'organisation narrative et du positionnement du narrataire.

Il est clair que l'anticipation (ou le retardement) induite par les coréférences rétrograde (ou différée) exercent une influence sur la façon dont l'audiospectateur perçoit les actions représentées visuellement, puisque, pour être comprises, elles demandent dans le premier cas un supplément verbal, alors que dans l'autre elles véhiculent une information narrative déjà connue. Ce second point est important dans la mesure où il détermine le régime de focalisation, voire le type de participation affective du spectateur au film: lorsqu'une voix-over annonce systématiquement et dans des termes fort précis ce qui va se passer par la suite, elle modifie les modalités de l'identification au personnage. Dans son étude des interventions auctoriales dans *Les Deux Anglaises et le continent* (François Truffaut, 1971), J. K. Sanaker note par exemple que le caractère «légèrement proleptique» de la voix provoque un désamorçage de l'émotion et du suspense<sup>682</sup>. Tout ce qui a été relevé dans cette section montre que le fonctionnement sémiotique de la coréférence verbo-icnique influe fortement sur notre «audiovision» du film.

## **VI.6 LA VOIX-OVER PARMIS LES «MECANISMES» DE L'ENONCIATION FILMIQUE**

On a vu lors de l'examen des présupposés attachés à la notion de «voix narrative» que son application au cinéma incite presque systématiquement les narratologues à prendre en compte la voix-over. Cette dernière fait office de «bon objet» dans la plupart des

681. Iouri Lotman, *Sémiotique et esthétique du cinéma*, Paris: Éditions sociales, 1977, pp. 31-32.

682. J. K. Sanaker, «La prose au cinéma ou Quand la voix off impose son rythme à l'image», in F. Vanoye (dir.), *Cinéma et littérature*, Paris: Université de Paris X, 1999, p. 103.

approches énonciatives, et cela même chez les auteurs qui dévient à ce type de considérations théoriques leur validité dans le champ de l'analyse filmique, à l'instar de J.-M. Schaeffer dans *Pourquoi la fiction?* Dans la section qu'il consacre au cinéma, Schaeffer développe une conception étroite de l'énonciation – selon lui, il ne peut être question de celle-ci qu'à condition que des actes de langage proprement dits soient effectués – et conclut que cette notion est inappropriée au cinéma. À son avis, un film est non seulement dépourvu d'énonciateur, mais il est également impossible de parler de « narration » filmique. Ce point de vue pourrait rejoindre celui que je soutiendrai dans ce chapitre (l'absence d'instance énonciative étant la meilleure prévention contre l'anthropomorphisme) si Schaeffer ne tombait pas dans une idéologie de la « transparence » qui, au même titre que les théories traditionnelles de l'énonciation filmique, fait l'impasse sur la dimension matérielle du film, et de façon flagrante sur sa dimension sonore (s'inspirant des *Essais sur la signification au cinéma* de Metz, il reconduit le primat quasi absolu de la bande-image). L'intérêt de son argumentation réside pour nous dans la restriction qu'il confère à sa thèse : « [...] à quelques exceptions près, le spectateur de cinéma ne voit pas le film comme quelque chose que quelqu'un lui montrerait, mais comme un flux perceptif qui serait le sien propre. »<sup>683</sup> On peut se demander ce qui est susceptible de « faire exception » ; en fait, l'auteur l'a signalé incidemment dans une partie antérieure de son ouvrage : « Sauf exceptions (voix-off), la parole au cinéma [...] permet à la fiction cinématographique de se passer de la narration au sens technique du terme. »<sup>684</sup> On retrouve donc, dans une théorie pourtant foncièrement opposée à la notion même de « narration filmique », l'affirmation du statut particulier de la voix-over. Selon Schaeffer, sa seule présence suffit à légitimer la prise en compte d'une instance énonciative. Par contre, pour les besoins de l'analyse, il exclut la voix-over de son étude (ainsi qu'en témoignent les parenthèses), alors qu'il serait peut-être productif de rapporter à celle-ci ce que la première citation impute à la seule image. L'idée qu'un spectateur puisse entendre la voix-over « comme un flux perceptif qui serait le sien propre » peut être associée à la théorie eikhenbaumienne du « discours intérieur ». Or il est clair que cette

683. J.-M. Schaeffer, *Pourquoi la fiction? op. cit.*, p. 301.

684. *Ibid.*, p. 257.

appropriation est enrayée par le fait que cette voix soit, de manière plus ou moins ostensible, le *produit* d'une énonciation.

Bien qu'il ne soit pas sémiologue, Schaeffer me paraît s'inscrire, en ce qui concerne les pages qu'il consacre au cinéma, dans un courant théorique que Dominique Chateau décrit ainsi :

« Dans la mesure où l'énonciation semble le plus souvent imposée de l'extérieur à l'image filmique (notamment par introduction du verbal), les sémiologues ont d'emblée incliné à étudier le récit filmique comme une sorte de « récit pur », comme une sorte d'histoire sans discours. »<sup>685</sup>

Une manifestation verbale peut être considérée comme « imposée de l'extérieur » lorsqu'il s'agit du boniment du cinéma des premiers temps, c'est-à-dire d'un énoncé linguistique qui s'ajoute à un « énoncé » filmique au stade de la projection (pour néanmoins produire une représentation audiovisuelle unitaire). Par contre, le verbal du cinéma parlant « semble imposer » une énonciation au film. La précision suggérée par Chateau me paraît importante, car l'énonciation verbale s'appréhende à mon sens en tant que phénomène de surface. Il s'agit avant tout d'estimer l'impression produite sur le destinataire du message.

Afin d'élaborer un modèle de l'énonciation filmique qui me paraît convenir à la prise en compte de la voix enregistrée, je propose de concevoir la notion d'énonciation, dans le domaine filmique, au sens tout à fait concret *d'ensemble des opérations humaines et techniques nécessaires à la production de la représentation*. Dans cette optique, on peut considérer que l'on se situe du côté de l'« histoire » au sens de Benveniste dès lors que le film procède implicitement à un déni des dites opérations. L'énonciation est par contre « discursive » lorsque les processus de production sont exhibés à l'intérieur même du texte filmique ou, si l'on veut englober également le dispositif, dans le contexte plus large de sa projection. J'entends donc la notion d'énonciation au « sens fort » du terme en postulant l'existence de « marques textuelles », bien qu'elles soient totalement différentes de celles que les poéticiens ont repérées dans le texte scriptural<sup>686</sup>. Cette distinction

685. D. Chateau, « Diégèse et énonciation », *art. cit.*, p. 122.

686. Chateau (*ibid.*, p. 121) distingue parmi les théories de l'« énonciation » celles qui confèrent à ce terme un sens faible (elles comprennent tout phénomène de communication) et celles qui l'entendent au sens fort (c'est-à-dire qui postulent la possibilité d'un repérage de marques énonciatives).

invalide la pertinence de l'étude des « déictiques », sans oblitérer pour autant leur valeur opératoire lorsqu'il s'agit d'examiner spécifiquement les manifestations verbales d'un film. En effet, le type de « renvoi » qui peut être « marqué » dans un film n'est pas relatif à une hypothétique « situation d'énonciation » au sens où l'ont conçue les linguistes (puisqu'ils axaient leur modèle sur la communication verbale), mais porte sur la matérialité du signifiant filmique : lorsque des sautes apparentes trahissent le fait que quelques photogrammes ont été soustraits de la bande pelliculaire, lorsqu'une poussière se balade de façon persistante sur la lentille de l'objectif du projecteur, lorsqu'un haut-parleur crépite parce que l'une de ses membranes est abîmée, l'existence matérielle du pôle machinique du dispositif cinématographique s'exhibe (ou est exhibée). Or, contrairement à la performance vivante qui est susceptible d'intégrer ces éléments parasites<sup>687</sup>, les conditions mécaniques de la production de la représentation, dénuées de capacités modulaires, ne font qu'en exhiber l'artificialité.

Cette matérialité ne s'affiche jamais pour elle-même – mon approche ne se veut pas immanentiste – mais doit être considérée telle qu'elle apparaît au spectateur, c'est-à-dire sur la base de ce qu'il infère des données présentes au niveau de la représentation. Ce marquage n'a donc d'existence que si le spectateur en prend conscience. Pour aborder cette dimension de la réception dans les limites de la sémio-pragmatique, on peut recourir à la notion de « savoir de l'arché » proposée par Jean-Marie Schaeffer dans *L'Image précaire*, où il discute la façon dont le spectateur « lit » une photographie. Dans le contexte qui est le nôtre, ce terme pourrait englober tout ce qui touche au savoir relatif à la fabrication des images et des sons. Schaeffer recourt à ce terme pour renvoyer à la matérialité du dispositif photographique au niveau de la genèse de l'image en postulant que, si l'observateur considère une photographie comme une empreinte, c'est sur la base de ses connaissances du fonctionnement de la technique utilisée :

« Comme l'interprétation des signes conventionnels, celle des signes naturels n'est possible que dans le contexte d'un certain savoir. Outre d'un savoir du monde, il faut encore disposer du

687. Voir P. Zumthor, « Le geste et la voix », *art. cit.*, p. 82.



savoir de l'arché: une photographie fonctionne comme image indicielle à condition qu'on sache qu'il s'agit d'une photographie et ce que ce fait implique.»<sup>688</sup>

L'optique de Schaeffer me paraît intéressante dans la mesure où elle déplace la problématique de l'indicialité de l'image dans un contexte pragmatique. Le principe qu'il examine sera considéré ici dans un sens élargi: le «savoir de l'arché» s'applique pour moi à la nature représentationnelle de toutes les composantes de l'expression cinématographique. Par conséquent, marquage énonciatif et inférences spectatorielles sont interdépendants: si un spectateur qui n'a jamais été confronté à l'acousmatisation technologique croit sincèrement, comme les prisonniers de la caverne platonicienne, que les figures de l'écran produisent effectivement les vibrations des masses d'air qui sont à la source des sons, alors les crépitements n'évoqueront pas pour lui la présence du haut-parleur.

Les quelques exemples de marquage cités ci-dessus (la poussière sur l'objectif de l'appareil, etc.) pourraient laisser croire que les conditions d'une énonciation marquée sont nécessairement liées à des dysfonctionnements techniques, et, de ce fait, échappent à toute «programmation» par ceux qui collaborent à la réalisation du film. D'aucuns diront qu'en ce cas, les marqueurs énonciatifs sont plus perçus qu'«énoncés». Un modèle qui se fonde sur la réception spectatorielle doit toutefois inclure de telles situations de réflexivité «involontaire», fût-ce au prix d'élargir le champ couvert par la notion d'«énonciation» (ce qui est de toute façon indispensable au cinéma, cette notion provenant de l'étude du langage verbal). En outre, les cas de figure évoqués ne concernent pas le texte filmique proprement dit, mais ses conditions de réalisation au sein du dispositif cinématographique. Il n'en va bien sûr pas toujours ainsi puisque, comme nous le verrons, les possibilités de marquage peuvent se situer à différents niveaux. Les éventuels «problèmes» techniques montrent néanmoins qu'il est possible d'évincer toute intervention d'une instance humaine, et qu'il est donc primordial de relativiser la place de cette dernière lorsqu'on échafaude une approche énonciative adaptée au médium. Certes, pour le spectateur, la voix introduit la composante humaine dans la machine; il me paraît toutefois méthodologiquement plus pertinent de s'en

688. Jean-Marie Schaeffer, *L'Image précaire*, Paris: Seuil, 1987, p. 41.

tenir, dans un premier temps, à la seule « machine ». Certaines pratiques cinématographiques ont d'ailleurs tenté de faire éclater la représentation en jouant sur l'exhibition de la machinerie, et cela également en ce qui concerne la composante sonore. À propos du *Traité de bave et d'éternité* (1951), premier film lettriste réalisé par Isidore Isou, Frédérique Devaux déclare :

« Outre qu'il ne s'agit plus de faire [...] un son net, propre, l'artiste accorde une valeur esthétique aux défauts de la mécanique d'enregistrement. À la limpidité du son, Isou préfère le brouillage qui participe au ciselant et permet de focaliser l'attention sur l'existence de l'outil de transcription des voix et des musiques. »<sup>689</sup>

Sans entrer dans le vocabulaire ésotérique et daté du poète lettriste (le cinéma « ciselant »), on constate que les « problèmes techniques » liés au son du film peuvent, dans le contexte institutionnel particulier d'une projection expérimentale, donner lieu chez le spectateur à ce que François Jost appelle une « présomption d'intention »<sup>690</sup> : les distorsions vocales ou les ruptures de synchronisme entre les paroles et l'image (le « montage discrédant ») seront rapportées à une intention communicative manifestée par le producteur du message. L'exemple du *Traité de bave et d'éternité* montre néanmoins qu'il est parfois difficile pour le spectateur d'établir s'il s'agit d'un effet concerté ou de l'influence fortuite de la technique ; le principe autographique de l'enregistrement sonore (le son vient s'imprimer de lui-même sur une surface inscriptible) s'accompagne souvent d'une ingérence humaine, et de toute manière ne conduit jamais à un duplicata exact de l'occurrence réelle. En fonction de l'horizon d'attente qui est le sien, le spectateur peut considérer un problème technique comme un effet de style (et inversement).

Les pratiques qui visent à exhiber le film dans sa matérialité ont parfois été élargies au dispositif même, conduisant les artistes à (ré)introduire la présence humaine et la fonction réflexive du verbal. Par exemple, lorsque Maurice Lemaître ou l'un de ses comparses se met lui-même en scène lors de la projection de l'un de ses

689. Frédérique Devaux, *Le Cinéma lettriste (1951-1991)*, Paris : Éditions Paris Expérimental, 1992, p. 49.

690. François Jost, *Un monde à notre image*, Paris : Méridiens Klincksieck, 1992, p. 85.

films ou, de façon plus radicale, à *la place de* la projection attendue<sup>691</sup>, nous retrouvons en quelque sorte un état antérieur de l'histoire du médium (le « cinéma parlé ») dans le contexte différent du « spectacle multi-média ». La parole, bien qu'elle ne soit pas produite mécaniquement, peut alors se référer au pôle machinique du dispositif.

Mon approche est donc fondamentalement différente de la plupart des conceptions énonciatives développées dans le cadre de la théorie du cinéma qui, à ma connaissance, n'intègrent jamais de façon systématique la composante de l'enregistrement sonore. Il me semble qu'au vu de l'artificialité sur laquelle repose la conjonction d'une voix et d'une image, la voix-over, pôle extrême de la déliaison vocale, ne peut être considérée comme « voix narrative » qu'au prix de sérieuses précisions liminaires et d'un modèle de l'énonciation filmique permettant d'envisager la représentation comme une partie intégrante du dispositif. C'est pourquoi j'écarterai les approches de l'énonciation au cinéma qui s'inspirent de celles qui ont été conçues pour le texte littéraire ou la communication orale, car elles ne sont exploitables qu'au niveau de leur principe général (le renvoi, adressé au lecteur via le verbal, à la source de ce dernier). Je me distancierai également des approches américaines du type de celle de David Bordwell, qui, dans une optique simplificatrice tendant à juger la théorie sur la base de sa seule productivité immédiate en termes d'analyse de film, refuse *de facto* le cadre des théories de l'énonciation. L'apport de ces théories me paraît indispensable, fût-ce au prix d'importantes adaptations. C'est pourquoi il est intéressant de revenir sur l'ultime ouvrage de Christian Metz, au titre fort significatif : « L'Énonciation *impersonnelle* ou le *site* du film. » Bien que cet essai s'appuie sur des principes passablement différents de ceux que je défends ici (notamment en ce qui concerne la nature de l'énonciation filmique que Metz juge intrinsèquement *narrative*), je trouve que l'auteur pointe avec justesse les dangers des perspectives théoriques qui mènent d'une part à une vision anthropomorphique inadaptée à la nature mécanique du médium cinéma, d'autre part à un plaquage artificiel des propositions de la linguistique. La remise en question du caractère opératoire des concepts de l'énonciation

691. Voir Dominique Noguez, *Éloge du cinéma expérimental*, Paris : Centre Georges Pompidou, 1979, pp. 102-104.

linguistique conduit Metz à prôner l'analyse de figures réflexives. Par conséquent, il se borne à repérer dans les films certaines manifestations de cette capacité qu'«ont beaucoup d'énoncés à se plisser par endroits, à apparaître ici ou là comme en relief, à se desquamier d'une fine pellicule d'eux-mêmes qui porte gravées quelques indications d'une *autre nature* (ou d'un autre niveau), concernant la production et non le produit, ou bien, si l'on préfère, engagées dans le produit par l'autre bout»<sup>692</sup>. Ce que le sémiologue exprime ici avec une grande finesse, c'est cette «impersonnalité» que j'attribuerai à l'origine machinique de la représentation filmique. Le postulat d'un refus de la «personne» me paraît nécessaire pour construire un modèle qui intègre facultativement le procédé de la voix-over car, comme le note Metz, «le spectateur attribue spontanément [...] les paroles de l'éventuel narrateur-off, ou du commentateur anonyme se voulant souverain, à un poste d'énonciation encore flou et incertain», celui-ci relevant dans tous les cas «d'une présence énonciative personnalisée»<sup>693</sup>.

Par rapport à l'hypothèse de Metz, je proposerai toutefois deux adaptations qui tiennent à l'objet particulier de mon étude (la voix) et à la relative déception ressentie à la lecture de son ouvrage lorsque, après cette introduction stimulante, nous passons à ce qu'il appelle la «visite guidée» des différents «paysages d'énonciation», termes aux connotations «naturalistes» pour le moins inappropriées à son approche. La prise en compte de la parole demande à intégrer, à l'intérieur de la conception générale de l'énonciation filmique, le système de l'énonciation linguistique que Metz écarte de façon radicale. En effet, comme on l'a vu en discutant certaines fonctions du bonimenteur, les déictiques verbaux peuvent être concrètement amenés à jouer un rôle dans l'organisation de la représentation. Il est en outre indispensable, pour éviter l'éparpillement des «paysages» metziens, d'avancer un modèle d'organisation énonciative permettant de systématiser les relations entre les différents paramètres du marquage (et non un certain nombre de constructions réflexives qui sont des actualisations spécifiques de ces paramètres), notamment ceux, fréquemment occultés par les théoriciens du cinéma, qui ont trait à la dimension sonore en général (et non seulement verbale). On fera

692. Ch. Metz, *L'Énonciation impersonnelle*, op. cit., p. 20.

693. *Ibid.*, p. 18.

intervenir la parole dans un deuxième temps, puisque je proposerais de greffer, selon des modalités qui restent à établir, le système énonciatif personnalisé de la tradition littéraire sur le modèle fondamentalement « impersonnel » de l'énonciation filmique.

Il est vrai que la conception *matérialiste* de l'énonciation filmique proposée ici n'est pas tellement en phase avec les courants dominants de la narratologie filmique, la tendance étant aujourd'hui à l'anthropomorphisation des instances énonciatives, comme on le constate notamment dans l'ouvrage *De la fiction* de Roger Odin, véritable somme des réflexions sur l'énonciation au cinéma et des thèses « sémio-pragmatiques » de son auteur. Odin précise d'ailleurs, dans l'introduction à l'ouvrage de Châteauvert sur la voix-over, qu'il « [...] s'inscrit dans un mouvement assez général de retour à l'humain [...] »<sup>694</sup>. Il me semble néanmoins que de nouvelles voies se profilent si l'on problématise les spécificités machiniques de la production sonore au cinéma. Je tiens néanmoins à conserver la notion d'« énonciation » pour deux raisons : d'abord, il me paraît important d'inscrire mon approche dans un courant théorique préexistant, la discussion de la pertinence des modèles s'avérant productive ; ensuite, la prise en compte d'éléments vocaux et verbaux demande à penser cette question en relation avec le sens courant de l'« énonciation ». François Jost considère cette acception du terme comme une aporie :

« Si le texte est une chose, il ne faut même plus parler d'énonciation. Ce concept n'a de sens que pour une pensée anthropomorphisante, pour autant qu'il renvoie à une instance humaine, située à l'extérieur du roman ou du film et responsable du discours narratif. »<sup>695</sup>

L'ouvrage de François Jost est précisément celui qui, dans le domaine francophone, a ouvert la brèche de l'humanisation (ou rappelé l'importance de cette dimension sur des bases plus rigoureuses), opposant la conception metzienne (dont il traite dans son premier chapitre intitulé « le film-machine ») au « récit autorisé » par l'« auteur », cette seconde perspective lui semblant plus adéquate. Jost est d'avis que le refus d'anthropomorphisation des

694. J. Châteauvert, *Des mots...*, *op. cit.*, p. 13.

695. F. Jost, *Un monde à notre image*, *op. cit.*, p. 31.

théoriciens de l'énonciation au cinéma est bien souvent paradoxal, l'abstraction de la notion d'«instance» se dissolvant dans la représentation concrète d'un individu: «[...] on doit bien avouer que le gain de désincarnation des instances est aussitôt perdu dans les pouvoirs que l'on est obligé de leur attribuer.»<sup>696</sup>

Toutefois, cette «désincarnation» s'avère doublement pertinente pour l'étude de la voix-over, tant au niveau du dispositif (elle est le produit d'une machine) qu'à celui de la représentation (elle ne s'incarne dans aucun locuteur). Alors qu'il s'agit pour Jost de retrouver l'humain par-delà la machine (notamment en étudiant le rôle de la voix), j'ai montré combien la nature foncièrement machinique du dispositif cinématographique avait des incidences décisives sur la réception spectatorielle. En fait, il faut à mon sens conférer à l'approche de Jost (prioritairement axée sur la dimension visuelle) les limites du «savoir de l'arché» du spectateur relatif à la technique sonore. Significativement, l'exemple à partir duquel il élabore son modèle énonciatif est un film muet du cinéma des premiers temps, *How it Feels to Be Run over* (Hepworth, 1900), dans lequel se pose avec acuité la question de l'ocularisation, alors que le bonimenteur n'est mentionné que pour être évacué<sup>697</sup>. Ainsi, on peut se demander si le «filmeur empirique»<sup>698</sup> dont parle Jost ne connaîtrait pas, dans le cinéma parlant, une instance équivalente pour la prise de son.

L'apport de l'ouvrage de Jost est néanmoins considérable, notamment en ce qui concerne le caractère polyphonique de l'énonciation filmique<sup>699</sup>. Cet emprunt à l'approche du linguiste Oswald Ducrot étant repris par Jean Châteauvert dans le cadre d'une analyse fouillée de la voix-over, je me propose d'interroger certaines hypothèses de Jost à travers ce qu'en fait Châteauvert, puisque l'objet de ce dernier est identique au mien. Pour motiver la perspective qu'il adopte, Châteauvert part du constat d'un «paradoxe théorique» sur lequel il revient fréquemment et dont il rend compte, dans son premier chapitre, de la façon suivante :

«Et, dès lors que l'on interroge la structure narrative mise en scène dans un film, dès lors qu'on essaie d'analyser la fonction d'un narrateur verbal qui se superpose à l'image et dit «Cette histoire a

696. *Ibid.*, p. 20.

697. *Ibid.*, p. 85.

698. *Ibid.*, p. 182.

699. *Ibid.*, pp. 78-79.

commencé il y a bien longtemps», on est inéluctablement conduit à un paradoxe théorique: d'une part on affirme que le narrateur verbal assume une responsabilité énonciative à l'égard de son seul discours verbal, mais d'autre part on souligne la propension de ce même narrateur verbal à segmenter le récit filmique et à induire des sous-segments «audio-visualisés». En d'autres termes, on «attribue» un segment de film tel un *flash-back* à un personnage tout en soulignant que sa responsabilité effective est circonscrite à son seul récit verbal – la responsabilité de l'image revenant au narrateur fondamental.»<sup>700</sup>

Pour résoudre ce paradoxe qui tient d'une part à ce que François Jost appelle la «paralepse constitutive du cinéma» («l'image en montre toujours indéfiniment plus que les mots qu'elle est censée visualiser»)<sup>701</sup>, d'autre part à un certain type d'organisation de l'énonciation (narrative), Châteauevert recourt au modèle de l'énonciation polyphonique d'Oswald Ducrot – très productif dans ce cadre – ainsi qu'au concept d'«attitude propositionnelle» qu'Umberto Eco développe dans sa théorie des mondes possibles, et qui a été préalablement proposé par le logicien Bertrand Russell. Or il me semble que si Châteauevert fait référence à ces deux approches, c'est notamment parce qu'elles lui permettent de faire du *personnage* le centre de gravité de sa réflexion. Le «paradoxe» qu'il évoque n'existe en effet que dans la mesure où l'analyste se place à l'intérieur de l'univers diégétique pour constater l'incapacité d'un locuteur à agir sur le montage du film. Toutefois, on peut concevoir cette situation aux plans strictement textuel et communicationnel comme une stratégie parmi d'autres de transmission d'informants au spectateur. Dans ce cas, les variations affectant leur distribution parmi différents canaux (notamment la voix-over et les images) ne surprend plus, à moins d'adhérer totalement, dans une perspective peu analytique, à la cohérence de la fiction. Les termes «narrateur» et «responsabilité» utilisés par Châteauevert font inévitablement reposer son approche sur des postulats anthropocentriques. Cette démarche ne me paraît justifiable que dans la mesure où, pour lui, celui qui «attribue un segment de film au personnage» n'est autre que le spectateur. Cette dimension pragmatique

700. J. Châteauevert, *Des mots...*, *op. cit.*, p. 21.

701. F. Jost, *Un monde à notre image*, *op. cit.*, p. 51.

incite alors le théoricien à privilégier la façon dont le spectateur échafaude une « image » de l'instance d'énonciation. Il est vrai que cette construction par l'audiospectateur d'une source du discours fait inmanquablement intervenir le comparant humain pour des raisons qui ont trait au confort que cette représentation procure face à « l'inquiétante étrangeté » suscitée par le dispositif. Il est cependant réducteur d'appréhender cette humanisation comme étant l'unique posture possible. Car l'approche de Châteauevert a pour corollaire d'évacuer toute dimension machinique. C'est pourquoi il n'applique pas sa conception « polyphonique » au signifiant de la représentation filmique. Au contraire, le modèle de Russell l'oriente vers l'humain plus que vers les autres pôles du dispositif, puisqu'il a été élaboré dans une perspective consistant à rendre compte de la « relation d'un sujet psychologique avec un objet extérieur »<sup>702</sup>. Or on constate que ce « sujet » n'est pas, pour Châteauevert, le spectateur en prise avec les conditions matérielles du dispositif, mais un *personnage* endossant le rôle de narrateur. À cet égard, l'attribution d'une attitude propositionnelle comme « /le personnage se souvient que.../ ou /le personnage rêve que.../ » est comparable aux « processus intérieurs » que Käte Hamburger considère comme des indices du statut fictionnel des textes. Elle postule en effet que la capacité de l'énonciateur à se situer au lieu de production de la pensée des protagonistes serait inconcevable s'il n'avait lui-même créé ces individus<sup>703</sup>. Une différence essentielle qui fait l'intérêt de l'approche de Châteauevert réside néanmoins dans le fait que l'attitude propositionnelle ne résulte pas d'une indication verbale explicite, et qu'elle peut donc être déterminée de multiples façons, par exemple grâce à un simple fondu enchaîné. Il n'en reste pas moins que la pertinence de cette théorie reste largement tributaire du statut fictionnel des êtres représentés.

Les considérations avancées par Châteauevert à propos du spectateur se limitent par conséquent à ce qui relève de la mise en scène d'une situation d'énonciation qui demeure interne à l'univers diégétique. L'humanisation induite par la voix et l'importante contribution des caractéristiques vocales dans le processus d'élaboration du personnage filmique exigent certes de discuter de tels phénomènes, mais il me semble plus pertinent, dans un premier temps,

702. J. Châteauevert, *Des mots...*, *op. cit.*, p. 69.

703. K. Hamburger, *op. cit.*, pp. 87-89.



de prendre la mesure des conséquences de l'impersonnalité foncière de l'énonciation filmique comparativement à ce que l'on trouve, par exemple, en littérature. Mon but n'est pas d'écarter la dimension pragmatique en traçant une ligne de fracture radicale entre réception spectatorielle et approche analytique, mais, au contraire, de prendre en compte le rapport du spectateur à la matérialité des signifiants visuel et phonographique. En ce qui concerne le son, les variations diachroniques rendent d'ailleurs l'exhibition de la matérialité technique tout à fait évidente: si, dans *On connaît la chanson*, Resnais insère «J'ai deux amours» chanté par Josephine Baker au lieu de faire réenregistrer le titre à l'aide de moyens techniques contemporains, c'est aussi pour que la déliaison (aspect central de l'esthétique de ce film) soit accentuée par le renvoi à la qualité de la fixation discographique de l'époque de la chanteuse.

Lorsque Châteauvert, de son côté, envisage la nécessité d'ajuster la théorie de Ducrot en tenant compte de ce qu'il appelle très justement la «matérialité du film», la «pluralité des matériaux qui composent le récit filmique»<sup>704</sup>, il aboutit au cas de l'assimilation, par le spectateur, de certains traits stylistiques à une figure d'auteur. Dès lors, les paramètres techniques susceptibles de faire office de marquages ne sont considérés qu'en regard de l'«appropriation du matériel expressif» par un cinéaste:

«Ainsi, lorsque je vois un film comme *E la nave va* (*Et vogue le navire*, Fellini, 1983), j'identifie certains éléments du film comme des marques d'énonciation du sujet-locuteur et de là je construis un sujet d'énonciation ( $\lambda$ ) à ce film; ce sujet, je ne l'appelle pas «film», mais bien «Fellini.»<sup>705</sup>

Il peut paraître quelque peu vain de mobiliser un tel appareil conceptuel pour faire *in fine* reposer l'argumentation sur cette «étiquette» qu'est le nom d'un cinéaste. Remarquons que cette position était déjà défendue par Jost, qui mentionnait dans sa typologie des énonciateurs le «supposé-réalisateur». Ce principe me paraît tout à fait pertinent: il se peut en effet que, par inférences intertextuelles, le spectateur rapproche un film donné d'autres réalisations

704. J. Châteauvert, *Des mots...*, *op. cit.*, p. 28.

705. *Ibid.*, p. 35.

attachées à la même «signature» sur la base de traits stylistiques communs. Jost se montre attentif au fait que la valeur énonciative «dépend très largement de ce que je sais sur le film, sur les conditions ou le lieu de production, sur le réalisateur, etc.»<sup>706</sup>.

Cette construction par le spectateur d'une figure auctoriale ne me paraît néanmoins s'effectuer de façon directe qu'en cas de reconnaissance de la voix ou de l'apparence du cinéaste. Si, par ailleurs, cette présence concrète de l'auteur est mise en scène en tant qu'instance narrative, comme la voix-over de Guitry dans *Le Roman d'un tricheur*, il est à mon avis tout à fait opportun de parler d'«énonciation cinématographique». Par contre, je trouve plus hasardeux de rapporter à ce type d'énonciation les «contre-plongées de Welles ou les multiples mouvements de caméra de Lelouch»<sup>707</sup> sans étayer, par une analyse du paratexte, ce qui oriente l'attention du spectateur sur de tels aspects stylistiques. Car il ne faut pas oublier que la grande majorité des productions cinématographiques, contrairement à l'exemple canonique des films de Welles (symptomatique de la rencontre toute naturelle mais parfois problématique entre la cinéphilie et la théorie du cinéma), n'a pas subi d'auteurisation. Châteauvert limite par contre avec raison la subjectivité de la perception spectatorielle en avançant la notion d'«usage» qu'il définit comme une «habitude collective de lire certains procédés comme référant à l'énonciation»<sup>708</sup>. À mon sens, l'ensemble des déterminations paratextuelles et institutionnelles qui pèsent sur la figure de «l'auteur» procèdent de l'*usage* au sens de Châteauvert.

En dépit de la nature sonore de son objet, Châteauvert ne souligne jamais l'importance que revêt sur un plan énonciatif la voix *enregistrée* du cinéma parlant, et ne considère pas la voix comme une composante *représentationnelle*. En fait, l'exacerbation du personnage dont témoigne sa théorie obéit au même déni des opérations matérielles de production du simulacre audiovisuel que l'esthétique de la «transparence» à l'œuvre dans la plupart des films. Pourtant, au cinéma, la voix n'est pas seulement un «discours verbal», mais un paramètre soumis à des impératifs de gestion technique, tant au niveau de l'enregistrement que de la synchronisation avec l'image. Une histoire des «machines parlantes» montrerait de façon éclairante à travers les commentaires émis

706. F. Jost, *Un monde à notre image*, op. cit., pp. 82 et 85.

707. *Ibid.*, p. 82.

708. J. Châteauvert, *Des mots...*, op. cit., p. 32.

dans les milieux de la critique, de la réalisation et des techniciens combien la perception de la restitution mécanique d'une occurrence vocale s'avère subjective et diachroniquement variable. Avant même toute appropriation de la technologie par un cinéaste, certains modèles de représentation sonore influencent la réception spectatorielle. Aujourd'hui, les standards labellisés comme le «son THX» (ou, à la fin des années 20, les appellations «Western Electric», «Tobis», etc.), mentions présentes le plus souvent dans la toute première image projetée, sont l'expression de ce réglage des attentes du spectateur par rapport à une certaine «fidélité» des sons. Dans un contexte institutionnel différent du mode de consommation dominant, par exemple celui du cinéma expérimental, il apparaît avec évidence qu'un travail sur la matière sonore proprement dite peut constituer l'une des «appropriations expressives» dont parle Châteauvert.

La référence que j'ai faite aux Lettristes ne doit pas laisser croire que j'envisage le traitement de la matière sonore strictement dans le cadre du cinéma dit «expérimental», car cette limitation conduirait à justifier, sur la base d'un partage net entre régime dominant et pratiques marginales, l'inutilité de cette conception pour des films narratifs. En effet, il arrive également que des productions hollywoodiennes courantes procèdent d'un travail de ce type. Ainsi, on peut certes analyser *Chaînes conjugales* comme une série de trois *flash-backs* dont la «responsabilité» énonciative incombe à trois personnages différents (selon l'attitude propositionnelle la plus courante /se souvenir de.../), la situation-cadre étant elle-même comprise dans un récit en voix-over d'un personnage hétérodiégétique qui se pose en créateur de monde. Il importe toutefois de noter les caractéristiques vocales des paroles sur lesquelles embraye l'audiovisualisation métadiégétique: il s'agit de la répétition sous une forme altérée d'une phrase préalablement présentée comme un monologue intérieur du personnage, ou adressée à celui-là par la narratrice *over*. Ce qui fait planer un climat étrange sur chaque *flash-back*, c'est le fait que s'opère, en son entame, une transition de l'humanisation induite par l'introspection du personnage vers la mécanique à laquelle renvoie la technique même de l'enregistrement. Pour produire cet effet, le technicien du son de *Chaînes conjugales* a recouru au Sonovox, un appareil conçu au début des années 40 pour accorder mécaniquement des bruits à la sonorité de paroles. Si la voix se fait bruit,

le contenu des paroles demeure néanmoins compréhensible. Par exemple, avant que Laura (Linda Darnell) ne se plonge dans ses souvenirs, le martèlement cadencé des gouttes d'eau qui tombent d'un lavabo dans un seau en fer s'amplifie (parallèlement à un travelling avant sur cette source sonore) puis se transforme en une phrase. C'est à l'issue de cette occurrence vocale mécanisée qu'un fondu nous fait passer à un autre lavabo tout aussi peu étanche, celui du modeste appartement de la mère de Laura. Ce procédé est utilisé pour le débrayage des trois analepses, toujours en lien avec une source sonore de type mécanique (le moteur à vapeur d'un steamer et le son d'une radio). L'altération déshumanisante des voix féminines introduit une angoisse qui tend à reverser sur l'univers diégétique le malaise éprouvé face à l'exhibition de la mécanicité de l'« espace de la production ».

*Chaînes conjugales* montre que l'examen de l'organisation discursive d'un film ne peut se passer, même dans une perspective strictement narratologique, d'une prise en compte des caractéristiques phonographiques de la voix. Si l'effet produit sur l'audiospectateur qui associe une voix à un personnage du film (par synchrèse, par inférences contextuelles ou narratives, par reconnaissance de la voix d'un acteur connu, etc.) ne doit pas être négligé, il importe toutefois de ne pas le confondre avec le niveau premier de l'organisation énonciative du film. C'est pourquoi je me propose d'établir, dans un premier temps, une taxinomie des marquages énonciatifs en tenant compte des conditions matérielles de la production du simulacre audiovisuel.

Par rapport au modèle déjà passablement « impersonnel » que j'ai proposé dans *La Fiction au cinéma*, il m'apparaît aujourd'hui préférable de supprimer le niveau de la « méga-énonciation » que j'avais emprunté partiellement à André Gaudreault. En effet, je crains, au vu de sa position surplombante et fédératrice, que ce niveau ne soit abusivement assimilé *de facto* à l'unité synchrétique d'un « narrateur » (significativement, Gaudreault parle de « méga-narrateur filmique »). Par ailleurs, il va de soi que mon objet demande ici un approfondissement de la question du marquage énonciatif relativement à la dimension sonore que j'avais passablement négligée. Dans le présent ouvrage, je me concentrerai donc principalement sur les sons<sup>709</sup>.

709. En ce qui concerne le marquage de la monstration et du montage de la bande-image, je renvoie au chapitre 3 de *La Fiction au cinéma*, où j'optais déjà pour une conception « impersonnelle » de l'énonciation filmique.

Afin d'isoler les composantes sonores en fonction de leur spécificité, je propose de recourir à la célèbre décomposition metzienne du langage cinématographique en «matières de l'expression filmique», qui présente l'avantage de tenir compte de la dimension «concrète» du support sémiotique. Il s'agit alors de retrouver un équivalent de l'origine énonciative que Kuroda décrit ainsi à propos du texte littéraire: «*Ce qui* (et non *celui qui*) crée la réalité fictive dans l'esprit du lecteur, ce sont les mots et les phrases que le lecteur lit, c'est-à-dire l'histoire elle-même et non l'auteur.»<sup>710</sup> À partir de catégories empruntées au linguiste Louis Hjelmslev («matière» s'oppose à «forme» et «substance»), Metz définit la notion de «matière de l'expression» comme «la nature matérielle (physique, sensorielle) du signifiant, ou plus exactement du <tissu> dans lequel sont découpés les signifiants»<sup>711</sup>. Le cinéma parlant résulte d'une combinaison de cinq matières de l'expression: la photographie mouvante, les mentions écrites, le son phonique, les bruits et le son musical. On peut donc concevoir les trois dernières comme autant de composantes susceptibles de faire l'objet d'un marquage de l'énonciation sonore. Celui-ci s'opère à l'un (ou plusieurs) des niveaux correspondant aux trois stades de l'élaboration de la piste-son que j'ai évoqués, en me référant à James Lastra, dans l'introduction de cet ouvrage: les niveaux prophonographique (concernant le lieu d'émission indépendamment de l'appareil d'enregistrement), phonographique (lié à la présence dudit appareil) et postphonographique (opérations effectuées ultérieurement à la prise de son). Rappelons que les sons filmiques sont, tout comme l'image, le produit d'une élaboration spécifique. La perception des sons d'un film se distingue en effet radicalement d'une situation d'écoute qui nous mettrait en contact de façon non médiatisée avec la source effective du son. L'écoute «première» a déjà été effectuée pour le spectateur par les différents intervenants du processus créateur de la piste-son (perchman, ingénieur-son, bruiteur, mixeur, etc.). Ne serait-ce qu'au niveau «prophonographique», que l'on a parfois tendance à oublier pour se concentrer sur le travail plus tardif des *sound designers*, de nombreux paramètres interviennent tels que l'acoustique du lieu de tournage – certaines configurations spatiales «profilmiques»

710. S.-Y. Kuroda, *art. cit.*, p. 283.

711. Ch. Metz, *Langage et cinéma, op. cit.*, p. 157.

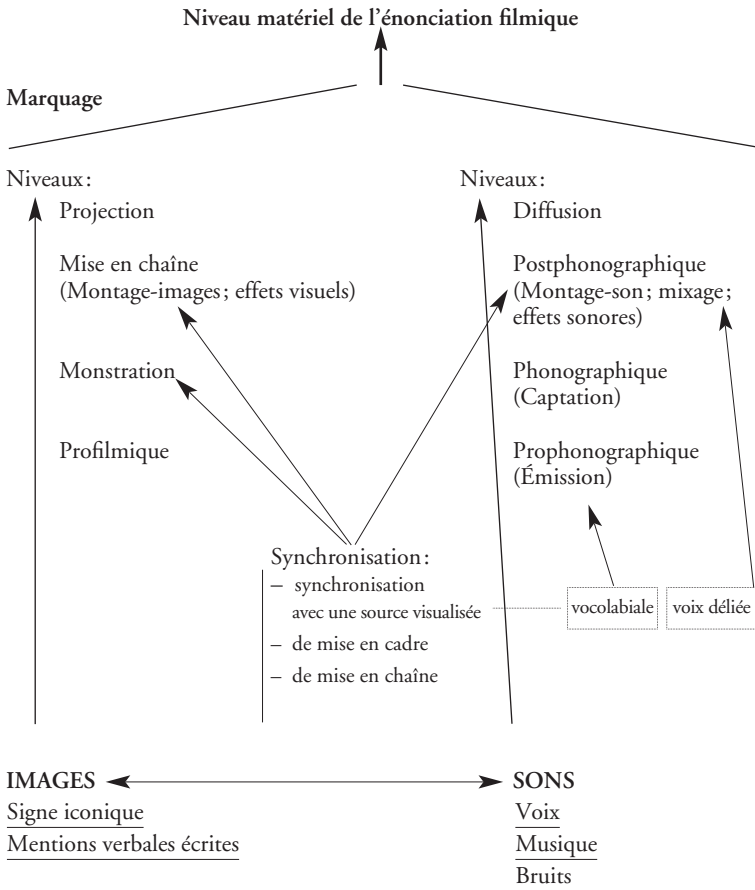
pouvant influencer le rayonnement de la voix – la portée de la voix des acteurs ou l'emplacement, l'orientation et les caractéristiques techniques des microphones.

Il faut en outre ajouter, dès lors que l'on s'interroge sur la perceptibilité de ces phénomènes par l'audiospectateur, la phase ultime du parcours du son: la diffusion en salle. Celle-ci engage à son tour des paramètres tant spaciaux que technologiques (enceintes, amplificateurs, etc.). Ce stade final est en fait déterminant pour tous les autres: les divers marquages n'ont pas d'existence en soi, mais dépendent d'une opération déductive du spectateur, auquel il incombe d'interpréter un trait sonore particulier comme un indice renvoyant au signifiant matériel de l'occurrence sonore. C'est pourquoi j'ai représenté, dans le schéma de la page suivante qui présente une conceptualisation provisoire de la question, la structure de l'énonciation filmique en figurant les inférences spectatorielles grâce à des flèches. Ce schéma se lit du bas en haut, c'est-à-dire de la dimension textuelle (les matières de l'expression) au marquage énonciatif inféré par l'audiospectateur. Dans la partie gauche, j'ai indiqué les équivalents visuels des mêmes niveaux en utilisant la terminologie filmologique ainsi que les catégories de Gaudreault (auxquelles j'ai ajouté la « projection » afin d'élargir la question au « dispositif cinéma » dans son entier). Cette représentation synoptique permet de faire en sorte que le son ne soit pas en reste par rapport aux concepts traditionnellement élaborés pour l'image.

En ce qui concerne le versant visuel de la représentation, rappelons qu'André Gaudreault a suggéré d'opérer une distinction entre (méga)monstrateur (profilmique: responsable de la mise en scène; filmographique: responsable de la mise en cadre) et narrateur (auquel incombe la mise en chaîne), ces deux instances étant subsumées par la notion de « méga-narrateur filmique »<sup>712</sup>. Si l'on soustrait à ce modèle sa composante anthropomorphique et la limitation au seul narratif, on obtient une distinction fondamentale entre trois niveaux: le profilmique, la monstration et l'opération de montage. Ces trois strates trouvent à mon sens une correspondance pour la « voix en tant que son » dans la tripartition *émission/captation/transformation*. La notion de « captation » permet de restreindre une première étape à la seule « prise de son »;

712. A. Gaudreault, *Du littéraire...*, *op. cit.*, p. 113.

## Les niveaux de marquage matériel de l'énonciation filmique



tout ce qui relève par contre de la création sonore appartient à la troisième phase que je qualifie globalement de « transformation », qui comprend le montage et le mixage son. Quant au pan sonore de l'instance, qui, pour Gaudreault, embrasse « mise sur film » et « mise en chaîne », il ne constitue pas dans notre optique une instance de premier niveau puisqu'il est intégré à un système plus vaste comprenant également les divers marquages énonciatifs affectant l'image. Ce problème ne se posait pas à Gaudreault dans la mesure où il faisait reposer son édifice théorique sur la seule composante visuelle. Cette occultation de la composante auditive s'explique dans *Du littéraire au filmique* par l'ancrage historique dans le cinéma des premiers temps (conçu comme silencieux, le bonimenteur n'étant pas encore aussi présent dans les préoccupations du théoricien qu'il le sera par la suite), ainsi que par la démarche explicitée dans le titre de l'ouvrage où le « littéraire » (récit scénique, et surtout *scriptural*) n'est pas envisagé sous l'angle de l'oralité. C'est pourquoi Gaudreault reprend le terme initialement proposé par Albert Laffay pour qualifier ce qu'il identifie comme une instance de premier niveau : le « grand imagier », c'est-à-dire une instance qui donne à voir, non à entendre (même si l'on peut parler également d'« images sonores »)<sup>713</sup>. Le marquage énonciatif doit être rapporté en dernière instance (mais non pas à une « instance ») au niveau de la représentation audiovisuelle dans son ensemble, voire à l'inscription de celle-ci dans le dispositif cinématographique. Dans mon schéma, je me suis contenté de faire figurer la représentation filmique, mais il est évident que celle-ci prend place dans le modèle tripolaire du dispositif (audiospectateur/machinerie/représentation) exposé au chapitre I.

La mise en parallèle des marquages visuel et sonore permet de souligner l'importance de leurs interactions. Par exemple, un usage « non transparent » de la synchronisation, notamment sous la forme de la déliaison vocale, peut conduire à mettre en évidence les opérations mécaniques qui président au couplage des deux composantes.

713. Notons que Laffay lui-même, comme on l'a déjà vu à propos de la voix-over, se montre plus attentif à la composante sonore que ne l'ont laissé entendre ceux qui se réfèrent à ses travaux. Ainsi développe-t-il l'idée d'une meilleure adéquation des sons avec l'image lorsque ceux-ci sont enregistrés en faisant l'hypothèse qu'ils partagent dès lors un même degré d'« irréalité » que la représentation visuelle : « C'est que la photographie, quel que soit son réalisme, ne laisse pas de transformer ce qu'elle produit. [...] *Il faut donc qu'il en soit de même des sons d'accompagnement* » (A. Laffay, *Logique du cinéma*, op. cit., p. 35). L'auteur précise d'ailleurs plus loin que « le cinéma prend la voix humaine et nous la restitue en chose » (*ibid.*, p. 49), témoignant ainsi d'une certaine conscience de l'existence de la matérialité sonore.



Cette question de la « transparence », fréquemment soulevée par les théoriciens du cinéma et la critique « idéologique » des années 70 en ce qui concerne l'image, peut surprendre lorsqu'on l'applique au son, puisque ce terme renvoie de façon littérale à l'aspect visuel. Alan Williams propose la notion d'« inaudibilité » (« *inaudibility* ») pour désigner un phénomène comparable à la transparence du montage classique, en postulant qu'« il existe, parmi les pratiques d'enregistrement sonore, l'équivalent exact de ce qu'on appelle une manipulation « invisible » (montage, etc.) sur la bande-image »<sup>714</sup>. Je préfère quant à moi conserver le terme de « transparence » pour parler du son (en le comprenant bien sûr dans un sens figuré), car l'élément pertinent pour le marquage ne réside pas dans le fait que l'on entende ou non quelque chose, mais dans la manière dont un certain son se fait entendre. Une notion identique permet en outre de souligner l'inévitable interaction qui s'opère entre les composantes visuelle et sonore qui, comme l'a noté Rick Altman, se soutiennent mutuellement dans l'entreprise de masquage de l'énonciation<sup>715</sup>.

Étant donné la faible perceptibilité du marquage des matières sonores comparativement à celui de l'image, Laurent Jullier conteste la pertinence d'une approche basée sur les inférences spectatoriennes comme celle que je prône ici. Il affirme en effet, à propos du rapport entre les éléments de la bande-son et l'audiospectateur, que « le processus de fabrication demeure inaccessible à l'auditeur »<sup>716</sup>. Selon lui, il n'est pas possible de remonter dans le sens inverse la chaîne de l'élaboration des sons (divisée dans son modèle en trois étapes : la fixation, les conversions et la diffusion) comme le fait le spectateur au cours de sa lecture de l'image. À mon sens, cette réserve est certes partiellement justifiée si l'on s'en tient à la comparaison avec l'image, mais il me semble abusif d'exclure foncièrement tout marquage de la matière sonore. Postulant que « le travail de la technologie ne peut jamais être totalement invisible », même si ce travail « vise à s'effacer lui-même », John Belton a par exemple montré de façon éloquente qu'il est possible de s'intéresser à l'« audibilité » de la piste-son proprement dite<sup>717</sup>. Dans

714. Alan Williams, « Is Sound Recording Like a Language? », *Yale French Studies*, N° 60, 1980, p. 61.

715. R. Altman, « Moving Lips: Cinema as Ventriloquism », *Yale French Studies*, N° 60, 1980, p. 69.

716. L. Jullier, *Les Sons...*, *op. cit.*, p. 34.

717. John Belton, « Technology and Aesthetics of Film Sound », in Elisabeth Weis et John Belton (dir.), *Film Sound. Theory and Practice*, New York: Columbia University Press, 1985, p. 63.

un ouvrage plus récent, Jullier considère néanmoins comme une preuve de l'imperceptibilité du signifiant sonore l'impossibilité dans laquelle se trouve l'audiospectateur de distinguer un son capté par une prise de son effective d'un son obtenu par synthèse. Toutefois, non seulement ce constat se rapporte uniquement à une période somme toute tardive de l'histoire du médium (on peut fort bien imaginer que cet argument soit prochainement valable pour l'image numérique), mais l'indistinction que Jullier évoque ne se produit que très rarement pour la voix, dont il traite somme toute assez peu. Les créations électroniques qui tentent de simuler des manifestations vocales ne se départissent d'ailleurs jamais totalement d'une certaine artificialité (qu'elles accentuent même volontairement, notamment pour provoquer une déliaison vocale motivée diégétiquement). De plus, même les films qui recourent massivement aux technologies de l'image de synthèse n'appliquent pas les mêmes principes de virtualisation à la voix. Par exemple, la première étape de l'élaboration de ce qui fut probablement le premier film à représentation réaliste intégralement réalisé en numérique, *Final Fantasy: Les Créatures de l'esprit* (Hironobu Sakaguchi, 2001), fut celle de l'enregistrement des dialogues par de vrais acteurs. Les images ne furent élaborées que dans un second temps, si bien qu'il a été possible de faire correspondre les mouvements aux voix. Il en va de même de tous les films d'animation récemment produits par Pixar ou Disney, dont la promotion exploite très fréquemment la mention des stars qui prêtent leur voix aux êtres infographiques.

Il est vrai que l'appartenance de la voix filmique au paradigme de l'humain a corrélativement pour conséquence de conduire les réalisateurs à souligner, dans ce qu'ils rendent « accessible » aux inférences spectatorielles, la *matérialité vocale* au détriment de la *matérialité de l'enregistrement* (ou *matérialité phonographique*). Je distinguerai donc deux types de « matérialité » : l'une est liée à la trace de l'événement unique et individualisé d'une émission présumée, l'autre aux conditions techniques d'existence du cinéma parlant. La matérialité vocale participe par ailleurs à des « procédures personnalisantes » qui résultent d'une construction particulière du discours filmique. Cette dernière privilégie généralement la diégétisation, c'est-à-dire le masquage plus que le marquage. Par contre, l'énonciation marquée au niveau matériel désactive, ne serait-ce que provisoirement, l'écoute causale et ses implications « natura-

listes» pour rappeler au spectateur que la voix entendue n'est qu'une reproduction.

Le marquage énonciatif ne nécessite pas forcément que l'audio-spectateur se fasse une idée précise des modalités de la fabrication des sons, mais simplement qu'il prenne globalement conscience de leur statut représentationnel. Le fait qu'il puisse être leurré sur le niveau de marquage n'est pas un critère qui invalide l'approche elle-même, car des phénomènes identiques peuvent être repérés pour l'image, notamment en ce qui concerne la monstration (on peut confondre un mouvement d'appareil effectué au tournage avec une opération de postproduction au banc-titre) ou le montage (par exemple, André Gaudreault n'a découvert que récemment la présence massive d'«arrêts-manivelle» dans les bandes des frères Lumière)<sup>718</sup>. Il est vrai néanmoins que la richesse figurative de l'image photographique constitue une prédisposition à la mise en évidence de la technologie visuelle: il est plus aisé de représenter visuellement l'appareil de prise de vues que de donner à entendre la technologie sonore en tant que telle, et non pour les sons qu'elle reproduit. C'est pourquoi une exhibition véritablement complète de la fabrication des sons exigerait l'intervention de l'image, ou du langage verbal dont le référent porterait sur la source effective des sons et sur leurs transformations aux différents niveaux phonographiques. On trouve ce type de remarques dans certains suppléments d'éditions DVD, où des explications données par des techniciens à propos de la réalisation d'effets sonores est simultanée à l'audition des sons diégétiques concernés à l'arrière-plan.

Dans *Monty Python, Sacré Graal (Monty Python and the Holy Grail, Monty Python, 1974)*, l'image se réfère explicitement à l'origine effective d'un son auquel l'audiospectateur attribue habituellement une source diégétique différente: les personnages de ce film sont à la fois partie prenante de l'action et assignés à la fonction de bruiteurs du galop de leurs propres chevaux (non visualisés), puisqu'ils se déplacent en entrechoquant des noix de coco (pratique traditionnellement utilisée en auditorium par les bruiteurs pour sonoriser la marche des chevaux). Dans le plan inaugural du film, l'illusion est momentanément entretenue du fait que

718. Voir notamment André Gaudreault et Philippe Marion, «Du sable dans les rouages du dispositif», in *Le Cinématographe, nouvelle technologie du XX<sup>e</sup> siècle, op. cit.*, pp. 229-242.

chevaliers et bruiteurs demeurent d'abord hors-champ, ce qui renforce l'effet de surprise provoqué par le détournement de la convention. On pourrait néanmoins imaginer, même si le cas est moins univoque, que l'artifice de la représentation sonore se dévoile à travers les caractéristiques de la seule bande-son en faisant entendre, sur une image d'écuycers chevauchant leur monture, les coques des noix tomber sur le sol de l'auditorium. Dans les films qui ne sont pas réalisés en prise de son directe, l'espace du studio d'enregistrement est en effet une « situation d'énonciation » première dont l'existence peut être plus ou moins masquée : les mauvais doublages antérieurs à l'apparition du son multipistes faisaient ressentir l'espace confiné de l'auditorium (niveau phonographique) qui contrastait parfois avec l'exigence, posée par l'image, d'une spatialisation des sons.

Il est évident que le degré de perceptibilité du marquage sonore est élevé lorsque ce dernier résulte des limites du système sonore. Les auteurs contemporains de techniques encore balbutiantes ont généralement tendance à relever ces phénomènes auxquels ils n'ont pas encore acclimaté leur perception. Un passage relatif à l'utilisation du phonographe par les pionniers du cinéma sonore, tiré de l'ouvrage *Le Cinéma* d'Ernest Coustet publié en 1921, est un exemple caractéristique de ce type de réception :

« Malheureusement, le phonographe n'enregistre pas seulement les paroles et la musique : les frottements du mécanisme, le grincement de la pointe qui a creusé sur le disque le sillon phonétique se traduisent par des vibrations parasites qui altèrent le son fondamental et ses harmoniques. »<sup>719</sup>

On peut certes considérer, à l'instar de Coustet, que le « bruit » lié au fonctionnement même de l'appareil constitue une source de désagréments. On peut également y voir la condition d'une utilisation artistique de la phonographie. Ainsi, dans un ouvrage datant du début de l'interrègne au cinéma (*Le Phonographe*, 1929), Cœuroy et Clarence appellent de leurs vœux la création d'« œuvres spécialement destinées à la machine parlante »<sup>720</sup>. En cela, ils inscrivent leurs considérations sur l'appareil phonographique dans l'un des

719. Cité par M. L'Herbier, *op. cit.*, p. 136.

720. André Cœuroy et G. Clarence, *Le Phonographe*, Paris : Éditions Kra, 1929, p. 52.

paradigmes importants de la réflexion des années 20 sur le cinéma, celui de la *spécificité* du médium. Il n'est dès lors pas étonnant que le terme de «phonogénie» ait cours à l'époque, le concept parallèle de «photogénie» ayant servi, à la suite de l'écrit de Louis Delluc (*Photogénie*, 1920), à désigner la capacité de la machine prothétique à révéler le visible sur un mode inaccessible à la perception quotidienne. La transposition de la photogénie dans le domaine du son ne s'est d'ailleurs pas limitée à la France, puisque l'année même de la publication du *Phonographe*, dont les auteurs encouragent le lecteur à cultiver la «déformation exquise» produite par l'appareil d'inscription/diffusion<sup>721</sup>, le Soviétique Adrian Piotrovski affirmait qu'il fallait «trouver une photogénie du son», «trouver les lois de la déformation sonore»<sup>722</sup>. Ce type de théories sur le son fait écho, dans le champ du cinéma, à l'argumentation d'un Iouri Tynianov ou d'un Rudolph Arnheim, pour qui les possibilités esthétiques sur lesquelles se fonde le statut artistique du médium naissent des différences qui séparent le film de la réalité profilmique<sup>723</sup>.

L'exhibition de la part machinique de la genèse des sons atteint son paroxysme lorsque apparaissent, grâce à l'électrification, des technologies capables de fabriquer des sons de toutes pièces ou d'altérer des sons préexistants par le procédé de l'«échantillonnage»<sup>724</sup>. Cette possibilité de se passer de la source sonore originale ou de la modifier jusqu'à la rendre méconnaissable, dont l'exploitation se trouve de nos jours encore facilitée par la numérisation, ne date pas des développements technologiques de la seconde moitié du siècle. Sous la forme non électronique du son optique, le procédé de la «synthèse» apparaît en effet au début des années 30 dans les travaux de cinéastes tels que Rudolf Pfenninger ou Oskar Fischinger. L'application de la technique «autographique» constitue même une des préoccupations de Moholy-Nagy dix ans auparavant. En 1923, cet artiste réalisait des rayures sur des disques afin, selon sa propre expression, «de faire du gramophone en tant qu'instrument de reproduction un instrument de production»<sup>725</sup>. Il s'agit là d'un procédé de «contact direct» identique

721. *Ibid.*, p. 129.

722. Cité par F. Albera, *Les Formalistes russes...*, *op. cit.*, p. 17.

723. Iouri Tynianov, «Les fondements du cinéma» [1927], in F. Albera, *Les Formalistes russes...*, *op. cit.*, pp. 73-99; Rudolph Arnheim, *Le Cinéma est un art*, Paris: L'Arche, 1989 [1933].

724. Voir L. Jullier, *Les Sons...*, *op. cit.*, p. 37.

725. Cité par Friedrich Kittler, *Film, Grammophon, Typewriter*, Berlin: Brinkmann & Bose, 1986, p. 74.

dans son principe à celui auquel Moholy-Nagy recourait pour ses «photogrammes», images obtenues sans appareil de prise de vues grâce à des objets disposés à même le support photosensible. En transposant ainsi une pratique «picturale» dans le domaine de la production sonore, Moholy-Nagy créait des sons qui n'avaient aucune existence acoustique préalable (aucune réalité phonographique), la matérialité du support devenant l'objet même de la représentation. Il entendait exploiter de la sorte, ainsi que le prôneront Cœuroy et Clarence, les possibilités propres du médium mécanique. Les nombreuses conceptions de l'autonomisation de la matérialité phonographique développées par certains artistes justifient que l'on ne considère pas uniquement les manifestations sonores de la technique elle-même comme de simples bruits parasites.

On rétorquera, dans l'optique de Jullier qui traite majoritairement de films réalisés à partir des années 70, que les imperfections techniques comme celles dont rendent compte certains commentateurs dans les années 20 n'existent plus aujourd'hui. D'un point de vue théorique, il ne faut cependant pas sous-estimer les implications de la nature *représentationnelle* des sons dans le cinéma parlant, qui se manifeste par exemple à travers une réduction ou une sélection des informations présentes dans la «réalité» reproduite (c'est-à-dire par rapport à la perception auditive habituelle que l'audiospectateur a du monde). Toute technologie sonore se définit par une certaine *bande passante* correspondant aux fréquences qui ne sont pas supprimées lors du passage de la captation à la restitution. Ainsi, un système aussi perfectionné que le Dolby Stéréo se caractérise par une bande passante comprise, dans le meilleur des cas, entre 50 Hz et 12000 Hz, alors que l'oreille humaine entend de 30 Hz à 15000 Hz<sup>726</sup>. C'est pourquoi, indépendamment de ce que le preneur de son choisit de retenir ou d'écarter, une certaine «perte» est inévitable.

La perceptibilité des altérations sonores dues à la technologie est certes faible, mais les concepteurs du film peuvent décider de les accentuer, par exemple pour donner l'impression qu'il s'agit d'un événement (visuel et sonore) «pris sur le vif» dont la captation a été aléatoire, ou tout simplement pour recréer une production sonore historiquement datée (voir la scène du phonographe dans *La Ligue des Gentlemen extraordinaires/The League of Extraordinary*

726. Selon J.-L. Passek, *op. cit.*, Paris: Larousse, 1995, p. 152.

*Gentlemen*, Stephen Norrington, 2003). Jacques Tati a montré qu'il était possible d'afficher le principe de sélectivité régissant l'obtention de sons filmiques, même si sa démarche radicale et anti-illusionniste est réappropriée pour partie au niveau diégétique, les caractéristiques de la « machine cinéma » étant transposées dans l'univers technocentriste représenté par *Mon oncle* (1958) ou *Playtime* (1967). Le renversement du marquage de la matérialité (sonore) sur le monde du film est un processus fréquent que je propose de qualifier « mise en phase », en référence à la notion avancée par Roger Odin, qui la définit comme « le processus qui (me) conduit à vibrer au rythme de ce que le film (me) donne à voir et à entendre »<sup>727</sup>. Afin d'éviter certaines imprécisions induites par cette définition<sup>728</sup>, je propose d'adapter quelque peu cette notion à notre contexte en l'envisageant dans un sens restreint. Il s'agira donc plus spécifiquement d'une stratégie consistant à tirer parti de l'effet d'inquiétante étrangeté suscité par la déliaison vocale (voir le chapitre I.1. Préalables théoriques : voix et locuteur dans les dispositifs audiovisuels) à des fins de renforcement de la participation spectatorielle, sans pour autant remettre en cause l'homogénéité et l'unicité de la représentation. Ainsi, les relations entre le spectateur et le signifiant filmique sont, ainsi que l'explique Odin, « homologues aux relations existant entre les éléments de la diégèse qui sont parties prenantes dans le déroulement de l'histoire »<sup>729</sup>. Avec le concept de « mise en phase », il est dès lors possible de rendre compte du reversement des effets produits par le dispositif à l'intérieur d'un seul de ses pôles, la représentation audiovisuelle.

Il me semble être du ressort d'une théorisation de l'énonciation filmique que de décrire les niveaux auxquels certains types de marquage de la matérialité sonore peuvent se manifester, comme cela a été fait pour le versant visuel et linguistique de la question. La voix du cinéma parlant est essentiellement une production machinique avant d'être le support d'un texte qui construit en creux la présence d'un sujet parlant. Dans la suite de cette section, je me propose donc de récapituler et d'illustrer les quatre niveaux de marquage sonore dont j'ai discuté les fondements. En inscrivant ces

727. R. Odin, *De la fiction*, op. cit., p. 38.

728. Voir Alain Boillat, « Jean-Marie Schaeffer, *Pourquoi la fiction?* et Roger Odin, *De la fiction* », *Iris*, N° 30, 2004, p. 165.

729. R. Odin, *De la fiction*, op. cit., p. 44.

remarques dans une conception graduelle du marquage énonciatif, on dispose d'un instrument de comparaison susceptible de rendre compte de pratiques filmiques diversifiées tout en permettant, si on élargissait l'approche à la dimension sociologique, d'intégrer la question du degré de la perceptibilité par le public des phénomènes acoustiques décrits. Cardinal affirme d'ailleurs dans sa réflexion sur la façon dont les théoriciens ont envisagé le caractère représentationnel des sons filmiques que l'«opposition entre représentation et reproduction ne nous semble [...] pas poser une différence de nature entre les deux termes, mais plutôt une différence de degré»<sup>730</sup>.

Le premier niveau de marquage, celui de l'*émission* vocale, est le seul qui soit commun au cinéma parlant et au cinéma parlé. Il ne requiert en effet rien d'autre qu'une source humaine, et en cela engage un degré de mécanicité largement inférieur à celui des autres niveaux (voire nul). L'exhibition porte plutôt sur la matière vocale: certains traits de la voix peuvent renvoyer au fait même de parler (raclements de gorge, signes d'épuisement, sautes de timbre, hésitations, etc.), de se mouvoir en parlant ou d'habiter un espace spécifique. Bien sûr, de telles marques, parce que considérées comme des «imperfections» par rapport à des normes de «neutralité», sont généralement supprimées au mixage. Il ne faut toutefois pas prendre la pratique dominante pour un absolu. On a vu par exemple combien la voix de l'écuyer de *Lola Montès* était marquée en termes de vocalité. C'est pourquoi, dans le schéma global des inférences spectatorielles induites par le marquage de l'énonciation filmique (voir le tableau récapitulatif à la dernière page de ce chapitre), j'ai orienté celles qui s'appuient sur la matérialité vocale vers le pôle de l'humain, non vers celui de la machine. Le niveau phonographique ne se réduit cependant pas à la production vocale proprement dite, mais comprend également le contexte spatial de cette production, c'est-à-dire notamment les caractéristiques acoustiques du lieu dans lequel la voix du locuteur est enregistrée. On sait que la généralisation du parlant a profondément influencé l'art des décorateurs qui durent tenir compte, pour la prise de son directe, du bruit «effectif» produit par les artefacts disposés sur le plateau. Leurs choix se répercutent sur le rendu de la voix des acteurs dans la mesure où certains matériaux de construction lui

730. S. Cardinal, *art. cit.*, p. 96.



confèrent une résonance particulière. Cet aspect se retrouve également à l'autre bout de la chaîne, dans la salle de cinéma, qui obéit à certaines normes de réverbération.

Au caractère humain de la manifestation vocale telle que la conçoit Paul Zumthor, il faut ajouter la présence graduelle de la mécanique que l'on observe dans les divers systèmes de sonorisation du cinéma. Même lorsqu'on la considère au niveau prophonographique (ou non-phonographique pour le cinéma parlé), la profération vocale peut être assistée d'un instrument technologique à la complexité variable (du haut-parleur à l'intervention de l'électronique). On peut constater alors un certain flou quant à la distinction entre les niveaux, mais il arrive que la situation soit désambiguïsée par la logique diégétique du film lorsque le représenté renvoie explicitement, comme cela est le cas dans les films non fictionnels, à une réalité prophonographique. Dans le cinéma documentaire, la différenciation entre émission et captation contribue en effet, lorsque l'occurrence sonore constitue un élément prégnant, à attester l'authenticité de la « source », ce terme étant à comprendre dans ses deux acceptions (en tant qu'origine physique et en tant que « document »). Dès lors, les altérations prophonographiques sont présentées au spectateur pour elles-mêmes. Elles peuvent résulter du seul comportement du locuteur qui travestit sa voix, comme dans *Hermann Slobbe, enfant aveugle (Blind Kid II, 1966, Johan Van der Keuken)*. L'enfant auquel ce documentaire est consacré est filmé alors qu'il assiste à une course automobile, mais le vrombissement des moteurs qui est synchronisé avec les plans sur les véhicules résulte, comme on nous le montre rétrospectivement, d'une imitation vocale réalisée par l'enfant. Le « leurre » initial du synchronisme (qui consiste à « subjectiviser » la représentation en nous faisant partager la perception intérieure d'Hermann à travers la façon dont il l'exprime par la voix), révélateur d'une intervention au niveau du montage-son, renforce la valeur d'authentification de la séquence insérée où nous voyons l'enfant devant un micro, utilisant parfois une boîte de conserve pour transformer sa voix. Indubitablement, cette performance de bruitage est à considérer comme relevant des niveaux prophonographique et phonographique (le rôle du microphone est souligné : Hermann le fait passer rapidement devant sa bouche pour simuler le déplacement d'un bolide), la piste-son du film se confondant avec l'enregistrement magnétique réalisé par l'enfant. Dans un film

de ce genre, il n'y a pas de raison de douter des capacités vocales de cet enfant aveugle qui rêve de devenir présentateur de radio.

D'autres altérations sont envisageables qui renvoient à un marquage de nature technologique, par exemple lorsque la voix est transmise à travers un système d'amplification électrique. Ce type d'appareil implique certes une captation-restitution du son, mais lorsqu'un film nous le donne comme participant de la réalité filmée, l'audiospectateur situera le marquage au niveau prophonographique. Il en va ainsi du documentaire de Donald Cumming significativement intitulé *Voice: off* (2003), dans lequel le cinéaste accompagne dans ses activités quotidiennes Gerald, un homme qui, suite à un cancer de la gorge, ne peut s'exprimer que par l'intermédiaire d'un appareil qui filtre et amplifie sa voix. Celle-ci transforme totalement sa voix, rendant son timbre métallique et les paroles parfois peu intelligibles. Ce n'est qu'une fois l'appareillage dont use le malade soigneusement montré à l'écran que l'audiospectateur est capable d'ancrer l'origine vocale dans le monde du film, de faire de cette voix *off* la voix *in* d'un *individu*. Les transformations prophonographiques de la voix – qu'il serait tout à fait incongru d'attribuer au travail du cinéaste sur la bande-son au vu du respect de la personne filmée qui transparait dans le film – ne cessent de renvoyer au laborieux processus de sa production. Il s'avère donc possible que la technologie du son modifie la voix à un niveau prophonographique, c'est-à-dire antérieur à l'enregistrement effectué par l'équipe du film. Le spectateur identifie ce niveau d'intervention sur des bases inférentielles qui relèvent également du savoir encyclopédique et intertextuel. Par exemple, dans les deux films de Van der Keuken et de Cumming cités, la lecture du spectateur est conditionnée par la reconnaissance du «genre» documentaire. Dans *Le Fantôme du Paradis* (*Phantom of the Paradise*, Brian de Palma, 1974), la voix du musicien défigurée par une presse à disques (William Finley) est également rendue par l'intermédiaire d'une prothèse phonique dont la technicité est exhibée. Toutefois, dans la fiction, l'origine des altérations sonores s'avance masquée (tout comme l'anti-héros du film): alors que l'enregistrement transforme la voix de l'acteur en la déshumanisant, les machines utilisées dans l'histoire visent à rétablir une sonorité humaine et l'intelligibilité des paroles. Dans l'une des séquences de ce film, c'est l'étape suivante du «parcours» technique de la voix qui connaît une mise en scène tragique: on y voit

le personnage méphistophélique de Swan manipuler les commandes d'une table de mixage afin de retrouver le timbre humain du «fantôme», ce dernier s'exprimant simultanément en face de lui par une série de crépitements qui, grâce à la médiation technologique, se font chant. En fait, deux niveaux sont simultanément convoqués dans l'activité de Swan: il s'agit d'une sorte de mixage réalisé lors de la prise de son elle-même, comme on peut effectuer un «montage à la prise de vues». Il en va de même dans *Sans soleil* (Chris Marker, 1983), où des sons synthétiques qui accompagnent par ailleurs le film en tant que musique extradiégétique sont montrés à travers le processus de leur création, Marker «documentant» le travail d'un artiste japonais.

Le premier de ces niveaux, la *captation*, peut s'envisager en termes d'énonciation, puisque la prise de son engendre parfois des manifestations sonores que l'audiospectateur rapporte au fonctionnement du matériel, à la présence même du technicien ou à ces deux aspects conjoints (perche que l'on déplace, manipulations sur un magnétophone, etc.). Il s'agit alors souvent de bruits considérés comme «parasites» qui ne se manifestent pas dans la voix reproduite elle-même, mais qui en influencent la réception. Comme émission et captation se complètent dans le processus de l'enregistrement, ces deux stades interagissent nécessairement dans les cas de marquage. Le film *C'est arrivé près de chez vous* (Remy Belvaux, 1992) offre, grâce au principe d'exhibition du tournage qu'il met en place, de nombreux passages marqués à tous les niveaux de l'énonciation audiovisuelle, si ce n'est que le montage est (presque) systématiquement nié, tant en ce qui concerne le son que l'image<sup>731</sup>. Le «film dans le film» imputé aux protagonistes de *C'est arrivé près de chez vous* passe pour être le film que l'on regarde: il s'agit d'un cas d'«identité» si l'on se réfère à l'échelle graduelle proposée par Lucien Dällenbach pour aborder des cas d'analogie entre l'œuvre «première» et sa mise en abyme<sup>732</sup>. Toutefois, le film de Belvaux constitue un cas-limite de redoublement de l'œuvre dans l'œuvre, puisque le paradoxe de l'identité ne peut pas être inféré de la structure du film (si ce n'est en ce qui concerne le statut énonciatif du générique de fin), mais uniquement de l'estimation du degré de réalité de ce qui est montré. Un paradoxe

731. À propos du marquage énonciatif de la composante visuelle dans ce film, voir A. Boillat, *La Fiction au cinéma, op. cit.*, pp. 111-114.

732. Lucien Dällenbach, *Le Récit spéculaire*, Paris: Seuil, 1977.

existe néanmoins au niveau de la logique narrative: le film ne clarifie pas pourquoi un criminel, certes exhibitionniste mais soucieux jusqu'à la paranoïa de ne pas faire parler de lui, accepte de paraître en flagrant délit dans un film. La diffusion du «film dans le film» dans le monde posé par *C'est arrivé près de chez vous* est tout simplement impossible, si ce n'est de manière posthume.

Le principal niveau de marquage sonore convoqué dans le film de Belvaux est celui de la captation. En effet, dans ce faux *reality show* où une équipe de cinéma suit un voleur et assassin dans ses activités de tous les jours, les trois «auteurs» présumés (cinéaste, caméraman et preneur de son) du «film dans le film» sont constamment présents, soit parce que les personnes filmées s'adressent à eux en direction du hors-champ (ou en regard-caméra), soit parce qu'ils apparaissent eux-mêmes à l'image. Le film se donne donc comme le résultat d'une prise (de vues et de sons) directe, gommant tout ce qui pourrait indiquer son statut fictionnel, même si le spectateur sait bien que les victimes ne sont pas mortes dans la réalité. Le caractère affiché de la dissociation entre l'image et le son, qui domine l'ensemble du film à travers l'individualisation des deux personnages censés être à l'origine de leur enregistrement, culmine dans une scène où tous les protagonistes sont à la recherche d'une gourmette perdue par l'assassin dans une usine désaffectée. Cette situation justifie que chaque personnage explore une autre parcelle d'espace tout en continuant à filmer (ou à enregistrer les sons ambiants), ce qui crée une distance physique entre le caméraman, situé aux côtés du voleur, et le preneur de son qui se trouve à l'étage inférieur. Dans un gros plan «*in*» sur le criminel dans lequel celui-ci s'entretient à haute voix avec le perchiste, la voix du premier est lointaine et affublée d'un important écho, alors que celle du perchiste, d'abord hors-champ, retentit très fort et distinctement. Ce renversement du rapport habituel entre l'intensité sonore et la distance séparant la caméra de la source des sons met en évidence les aléas qui, dans cette scène, touchent la captation. Les difficultés de la prise de son s'aggravent dans le plan suivant, lorsque le voleur s'adresse en regard-caméra au filmeur. En effet, ses lèvres remuent, mais au lieu de sa voix nous n'entendons rien d'autre que des bruits de pas sur une surface métallique, jusqu'à ce que le preneur de son apparaisse à l'arrière-plan. Avec l'avancée progressive du *perchman* en direction du locuteur, la voix du criminel se fait de plus en plus audible. Ce cas de déliaison par

suppression, puis atténuation de la voix est sans conteste rapporté par l'audiospectateur à la « situation d'énonciation » sonore. Plus tard, le preneur de son se fait abattre par un tireur embusqué. À partir de cet instant, le son se résume à un bruit assourdissant avant de disparaître complètement. Les variations qui touchent la qualité et la présence du son en fonction des situations narratives dans lesquelles le preneur de son est impliqué constituent un marquage sonore au niveau de la captation. Bien sûr, cette captation *in situ* est mise en scène. Cette facticité est par exemple révélée par la présence irrégulière du bruit de la caméra, qui disparaît soudainement dans les moments forts du point de vue dramatique. *C'est arrivé près de chez vous* ne procède pas d'une démarche anti-illusionniste, mais, au contraire, d'une volonté d'accentuer l'effet de réel par une esthétique pseudo-documentaire. C'est pourquoi la référence indirecte au bruit du projecteur, aspect dont on a vu qu'il a été commenté par nombre de théoriciens et historiens du cinéma, disparaît dès lors que le bruit ne renvoie pas explicitement au filmeur diégétique. La réflexivité permanente étant délibérée et construite, rien ne nous assure que, dans les faits, le marquage ait effectivement été réalisé lors de la prise de son (certains sons ont d'ailleurs certainement été créés ou modifiés en studio). Ce qui importe, c'est le statut énonciatif du film tel qu'il s'exhibe au spectateur.

La façon dont Christian Metz interprète les catégories de Benveniste a profondément influencé la perspective généralement adoptée dans le champ du cinéma. En affirmant que le « discours se déguise en histoire », Metz semble inclure l'idée d'un rapprochement entre *discours* et « pureté » (forme première, non « déguisée ») d'une part, entre *histoire* et manipulation (illusion du « déguisement ») de l'autre, alors qu'il ne s'agit là en fait que de deux modalités énonciatives opposées qui, l'une comme l'autre, résulte d'une stratégie communicationnelle spécifique<sup>733</sup>. Les théoriciens « matérialistes » qui ont développé le même type de réflexion en exigeant d'un film politique une remise en cause de l'idéologie de la représentation n'ont fait que renforcer une interprétation en termes de « vérité » ou d'illusion, ce à quoi un système énonciatif donné ne peut renvoyer en propre. Je pense notamment à la série d'articles publiés par Jean-Louis Comolli dans les *Cahiers du Cinéma* en 1970-1971 : à ceux qui prétendent par exemple que le film *Othon* (Straub et Huillet,

733. Ch. Metz, *Le Signifiant imaginaire*, op. cit., p. 113.

1969) est « inintelligible, ‹inaudible› tel que la difficulté d'écoute et donc de lecture grève, gêne considérablement la réception de tout message politique » et valorisent *L'Aveu* de Costa-Gavras (1970), Comolli répond que *L'Aveu* manque « de toute mise en question des conditions de production/écriture/diffusion/lecture du film. »<sup>734</sup> *C'est arrivé près de chez vous*, film éminemment discutable d'un point de vue éthique, illustre combien le régime du « discours » n'est pas incompatible avec l'illusion, et peut même décupler l'effet de tromperie. Le film de Belvaux montre également que le marquage au niveau de la captation des sons s'établit bien souvent sur la base d'une situation d'énonciation exposée visuellement au spectateur. Il paraît donc indispensable de considérer le marquage énonciatif comme un effet combiné des composantes visuelles et auditives. Pour cette raison, il est peu fréquent de rencontrer des voix-over dont la matérialité sonore de l'enregistrement renvoie sans ambiguïté aux conditions de sa captation. Si l'usage traditionnel de la voix-over (dont la « transparence » est à l'origine de son association à la « voix narrative ») semble exclure ce type de marquage, on peut trouver cette dominante *discursive* dans des films non fictionnels dont la bande-son est le résultat d'un enregistrement effectué sur les lieux mêmes que l'image visualise et n'entretient avec les locuteurs de l'écran qu'un synchronisme très large (comme dans la plupart des films de Jean Rouch ou dans *Unsere Afrikareise* de Peter Kubelka). Les conditions de tournage précaires de ce type de films provoquent par ailleurs des imperfections dans la prise de son directe qui sont perceptibles par l'audio-spectateur.

La *transformation* de la voix d'un locuteur et l'intervention ostensible du montage-son, deux opérations qui constituent les principales modalités de marquage postphonographique, apparaissent dans la plupart des déliaisons vocales qui exhibent la technologie du son. On peut citer, parmi les nombreux exemples de déliaison que l'on trouve fréquemment dans le cinéma fantastique, le film *L'Exorciste* (*The Exorcist*, William Friedkin, 1973), qui est à cet égard particulièrement intéressant parce que les nouveaux standards qu'il a posés dans le genre de l'horreur impliquent un travail important sur la piste-son, ce qui a été reconnu par la profession

734. Jean-Louis Comolli, « Film/politique (2). *L'Aveu*: 15 propositions », *Cahiers du Cinéma*, N° 224, 1970, p. 48.

puisqu'il a remporté l'oscar pour cet aspect technique. Les déliaisons s'effectuent par rapport à la voix de la jeune fille Regan (Linda Blair), doublée par Mercedes Mc Cambridge et transformée au mixage. En dépit de passages *gore*, les principaux effets horrifiques du film de Friedkin reposent sur le surgissement de la voix du Malin qui vocifère des obscénités à travers la bouche de la fillette. Contrairement à des situations de suspense où le spectateur est amené à s'interroger sur l'origine spatiale de la voix, dans *L'Exorciste* on sait toujours d'où le diable s'exprime, puisqu'il a fait sien le corps de Regan pour communiquer avec l'entourage de celle-ci. C'est justement cette absence d'ambiguïté quant à l'ancrage spatial de la voix qui est déstabilisant, puisqu'elle apparaît foncièrement inadaptée à son lieu d'émission. La déliaison vocale y est d'autant plus prégnante que le corps de Regan est l'élément nodal du récit. Le lien entre le corps et la voix dépasse le seul synchronisme des mouvements labiaux, car la possession, manifeste par la voix de l'Autre, s'accompagne d'une mise à mal de l'intégrité physique (la jeune fille enjoint les médecins à la violer, se lacère le visage, se mutile les parties génitales, effectue des mouvements improbables, prend des postures contre-nature, etc.). Dans cet anéantissement s'opère la dénégation du corps (sexué) en tant que lieu du sujet (parlant). Alors que la voix cavernueuse est connotée comme masculine, la jeune fille subit dans sa sexualité même les attaques répétées de l'être qui habite son corps. Cette présence de la corporéité profilmique induit par contre un effacement de la composante technique du dispositif. La *mise en phase* marche ici à son plein : la voix constitue la préoccupation principale des protagonistes, notamment lorsque le prêtre Karras effectue un enregistrement des propos de Regan sur bande magnétique afin d'obtenir de l'église l'autorisation d'entreprendre l'exorcisme, et constate que certaines phrases sont dites à l'envers. Dans ce type de situation, le film oscille entre marquage et masquage : si le procédé technique dont les ingénieurs du son ont usé est exhibé dans la diégèse, c'est en fait pour être mieux occulté, car la voix monstrueuse, souvent mêlée à des grognements, renvoie à une dimension bestiale et aucunement mécanique. Par ailleurs, l'enlaidissement progressif du visage de Regan conduit à atténuer l'effet de la déliaison puisque la transformation de l'apparence de l'actrice se fait l'écho des altérations opérées sur la bande-son. Par contre, deux autres voix émanant de la bouche de la fillette (dans la version *director's cut* du film), celle de

la mère de Karras et celle d'un mendiant, relancent l'exploitation du potentiel d'étrangeté de la déliaison vocale et peuvent être interprétées comme des cas de marquage postphonographique. Lorsque Regan s'écrie *off* «aidez un ancien enfant de chœur», il s'agit d'une réplique et d'une voix déjà entendues auparavant dans la bouche d'un clochard, si bien qu'on a l'impression d'assister à un équivalent sonore de ce que Metz a appelé pour l'image un «insert diégétique déplacé»<sup>735</sup>. Aussi *L'Exorciste* exploite-t-il l'impact de la dissociation du corps et de la voix, qui renvoie, au niveau du support de la représentation, à l'existence d'une différence matérielle entre la piste-son et la bande-image. Toutefois, un processus de diégétisation vient occulter la réflexivité induite par la mise en évidence de la matérialité phonographique des sons: cette dissociation se voit intégrée dans la diégèse sous la forme de l'opposition traditionnelle du Bien et du Mal. Le marquage de la matérialité phonographique est donc passablement atténué par une motivation diégétique de la nature «construite» de la voix.

La référence à la machinerie est par contre plus importante lorsqu'un film engage la représentation d'une origine mécanique de la voix. Comme tous les paramètres évoqués ici, la transformation doit être conçue de manière graduelle. En effet, il existe de nombreuses altérations qui ne présentent aucune pertinence perceptive, la voix d'un individu n'étant d'ailleurs pas toujours exactement identique dans tous les espaces et à tout moment de sa vie. De plus, l'auditeur ne dispose pas d'«échantillons» de comparaison, si ce n'est mentalement, en fonction de sa mémoire auditive qui peut être plus ou moins bonne, comme on est plus ou moins physionomiste. C'est pourquoi la transformation n'est évidente qu'en cas d'utilisation de certains procédés techniques (filtrage, chambre d'écho, etc.), notamment lorsque le film déshumanise une voix humaine. Ce type de voix se rencontre fréquemment dans les films de science-fiction dont l'univers comprend des êtres cybernétiques.

Cependant, là aussi, le marquage énonciatif n'est généralement pas de mise, l'irréalité du monde représenté étant compensée par une forte immersion du spectateur. La transformation de la voix y est motivée *diégétiquement* par le statut non humain du personnage et *visuellement* par un phénomène de synchrèse qui pousse

735. Ch. Metz, *Essais sur la signification au cinéma*, op. cit., p. 126.



l'audiospectateur à attribuer automatiquement la voix entendue à l'individu présent simultanément à l'image. Ces conditions tendent donc à exclure le statut *over*, bien qu'il arrive que l'ancrage visuel ne s'effectue pas. Un cas célèbre d'absence (momentanée) d'ancrage précis est celui de la voix de l'ordinateur parlant HAL de *2001, l'Odyssée de l'espace* (Kubrick, 1968) dont traite Michel Chion dans le cadre de son analyse des «acousmètres»<sup>736</sup>. Un exemple analogue me semble particulièrement intéressant : il s'agit de la voix attribuée au pilotage automatique du vaisseau Nostromo dans *Alien, le huitième passager* (Ridley Scott, 1978). L'acousmètre du Nostromo se caractérise par une voix portant les traces de fortes distorsions, mais qui semble plutôt féminine, les personnages appelant significativement cette machine «Mother». D'abord associée, dans la version doublée en français, à un écran d'ordinateur (qui joue le même rôle d'ancrage vague, inapte à la synchronisation stricte, que «l'œil» rouge dans *2001, l'Odyssée de l'espace*) et non à une «source sonore manifeste»<sup>737</sup>, cette voix est ensuite produite (cette fois également dans la version originale) à l'occasion du déclenchement de l'autodestruction du vaisseau spatial. Elle retentit alors dans l'ensemble de ce huis clos «utérin», les décors jouant d'ailleurs sur une possible confusion entre le mécanique et l'organique. Voix *off* plus que *over* du fait de son inscription dans un espace diégétique partiellement visualisé, la voix maternelle du Nostromo est marquante (et *marquée* à un niveau postphonographique) parce qu'elle se présente comme l'altération délibérée d'une voix humaine, faisant ainsi écho aux diverses voix modifiées

736. Michel Chion, *La Voix au cinéma*, Paris : Éditions des Cahiers du cinéma, 1982, pp. 34-35. Chion qualifie d'«acousmètres» des personnages – et donc des supports individués de la voix – que l'on entend sans qu'ils soient visibles à l'écran, tout en étant susceptibles d'apparaître à tout moment, subissant alors une «désacousmatisation». Sa thèse consiste à postuler un imaginaire de la voix qui résulte d'une corrélation entre cette situation acousmatique et certains attributs fictionnels, tous liés à une position de pouvoir : l'acousmètre s'avère ubiquitaire, omniscient, omnipotent et dispose d'une vision panoptique sur le déroulement des événements. En effet, en n'étant pas soumis à un ancrage donné dans l'espace diégétique, il acquiert une totale autonomie par rapport aux contraintes de vraisemblance imposées par une gestion réaliste du «point de vue» au sens large, qui comprend à la fois ce qui est vu, entendu et su. Or, contrairement au locuteur *over*, il peut en faire un usage immédiat dans l'univers diégétique. Même si la voix n'est pas véritablement détachée du corps (selon la logique posée dans le monde du film), elle bénéficie sur le plan fantasmatique des conséquences de cette désincarnation, fissent-elles provisoires. On pourrait dire que le narrateur *over* (homodiégétique) dispose d'autant plus fortement des pouvoirs conférés par Chion à l'acousmètre qu'il peut n'être jamais contraint à une désacousmatisation.

737. « Une source sonore est manifeste lorsqu'elle manifeste par un mouvement visible dans l'image son statut de source en train d'émettre » (R. Odin, *Cinéma...*, *op. cit.*, p. 254).

des personnages que le film motive diégétiquement grâce au contexte technologique (retransmission hertzienne, scaphandres, répercussions des sons dans les bouches d'aération, etc.). Cette mécanisation de l'univers du film introduit un transfert du marquage vers une origine diégétique (fictive). Néanmoins, il est un cas particulier de « citation filmique » où un avatar de la voix de Mother apparaît plus ostensiblement comme le résultat d'une opération technique, effectuée en l'occurrence au niveau du montage-son : il s'agit de la « greffe » de cette voix semi-robotique dans le film *Alien 3* (David Fincher, 1992). Ce deuxième *sequel* du film de Scott présente un redoublement de l'acousmatisation de la voix féminine d'un enregistreur de vol (celui du vaisseau « Sulaco » avec lequel l'héroïne s'échappe à la fin de *Aliens*, le second épisode réalisé par James Cameron) grâce à une source seconde qui en médiatise la transmission. Lorsque la tête arrachée et abîmée d'un *cyborg* – dont on discerne encore une partie de la bouche, garante du synchronisme vocolabial – est branchée sur la boîte noire du « Sulaco », la face de cet androïde « cite » une phrase qui a été prononcée par la voix de l'ordinateur du vaisseau disparu (identique à celle de « Mother »). À ce moment-là, l'androïde abandonne sa voix faiblarde (doublement transformée, pour signifier d'une part l'aspect électronique, d'autre part l'agonie de la machine) et endosse la voix du « Sulaco ». Or ces paroles (une alerte incendie) ont déjà été entendues par le spectateur durant le générique d'*Alien 3*, où le statut de la voix était intermédiaire entre le « *off* » (pour un spectateur familiarisé avec l'univers posé dans les épisodes antérieurs) et le « *over* » (la fragmentation du visible en flashes très brefs ne permettant pas de statuer sur l'éventualité d'un ancrage dans la diégèse). Située au tout début du film, cette voix emprunte à l'usage courant de la voix-over sa fonction d'inauguration du récit. Lorsqu'elle est littéralement (et « matériellement ») citée par le robot, elle fait par contre ostensiblement ressentir l'existence d'un « montage-son », motivé diégétiquement par les compétences imputées au *cyborg*. Les derniers plans d'*Alien 3* poursuivent cette exhibition de la matérialité phonographique en reprenant exactement les phrases que Ripley avait enregistrées à la fin du premier épisode. Cette greffe sonore issue du film de Ridley Scott ne renvoie pas seulement aux possibilités du montage-son, mais aussi à celles du mixage, puisqu'on nous fait entendre une voix très altérée par la dégradation supposée de l'enregistrement.

Cette altération d'une occurrence vocale entendue précédemment clôt le film de façon symbolique (la dégradation de la bande rappelle la mort de l'héroïne, le dernier mot cité étant « terminé ») et provoque un effet d'annonce de l'épisode suivant (la répétition des paroles de Ripley – ou *re-play* – sous une forme transformée techniquement préfigure le thème du clonage d'*Alien 4, la résurrection*). *Alien, le huitième passager* et son second *sequel* illustrent les possibilités d'un travail sur la matière de l'enregistrement vocal qui s'affiche avec un degré de marquage considérable sans pour autant contrevenir aux exigences de la cohérence diégético-narrative<sup>738</sup>. Nous sommes donc loin de la démarche lettriste consistant à rompre l'illusion audiovisuelle, mais le traitement de la thématique « mécaniciste » induit, me semble-t-il, une forme partielle de réflexivité.

Un marquage postphonographique par exhibition du montage-son s'effectue chaque fois que s'opère une « citation » dont l'audiospectateur est susceptible d'identifier l'origine (à l'instar de certains *samples* vocaux introduits dans les titres de musique rap ou techno). Cette pratique de la citation sonore est poussée à l'extrême dans *Passage à l'acte* (1993) du cinéaste autrichien Martin Arnold. Ce film se présente sans équivoque comme le recyclage d'un film parlant antérieur (*Du silence et des ombres/To Kill a Mockingbird*, Robert Mulligan, 1962) sur lequel il opère une dilatation par répétition de photogrammes et des sons correspondants. Ainsi les trente-trois secondes de la séquence originelle, qui représente une discussion au sein d'une famille, deviennent douze minutes, la synchronisation vocolabiale étant exhibée de façon quasi hallucinogène (la répétition porte parfois sur un seul phonème) : les voix, dont les paroles sont inintelligibles, ne sont plus que rythme. Le film d'Arnold provoque chez l'audiospectateur une prise de conscience des manipulations qui ont été pratiquées sur le son à un niveau postphonographique, puisqu'il s'agit de la reprise d'une production audiovisuelle antérieure.

En ce qui concerne le dernier niveau, celui de la *diffusion*, il ne m'est bien sûr pas possible de citer un film en particulier (à moins

738. On trouve un processus semblable au niveau de l'image dans *Alien 3* lorsque Ripley, décrivant la première apparition du monstre, remarque qu'il se déplace différemment de ceux qu'elle a connus auparavant. L'histoire technologique du médium – le film de Fincher dispose désormais de la technologie numérique pour animer l'*alien* – est transférée de façon manifeste dans le monde du film.

d'élargir au terme « film » des performances comme celles des Lettristes), puisque ce marquage s'effectue en fonction des conditions mêmes de la projection, par définition labiles. Tout ce qui a été dit du « son comme attraction » lorsque j'ai abordé dans une perspective historique l'exploitation (pré)cinématographique participe d'un marquage de la diffusion. Ce type de réception acoustique implique l'activation de ce que Laurent Jullier appelle « l'écoute causale de premier niveau » durant laquelle l'auditeur appréhende le haut-parleur « non pas comme un médiateur, mais comme source naturelle active »<sup>739</sup>. Il en va de même d'éventuelles autres composantes de la machinerie de restitution sonore (comme, à l'époque du « muet », les systèmes semi-mécaniques de synchronisation de Delacommune, Beck, Lachmann, etc.), du bruit provoqué par le projecteur ou des claquements dus aux poussières ou rayures sur la pellicule. Cette étape de la diffusion peut néanmoins être mise en scène dans le film, comme cela est systématiquement le cas avec les voix que l'on pourrait appeler à la suite de Michel Chion *on the air*, car elles sont marquées par les caractéristiques acoustiques de leur espace de réception<sup>740</sup>. Dans le générique du récent *Territoire des morts* (*Land of the Dead*, George A. Romero, 2005), un malaise renvoyant au caractère post-apocalyptique de l'univers représenté est créé par des voix « over » de commentateurs radio avertissant la population des risques d'épidémie. Le brouillage de cette transmission, auquel correspondent, sur le plan visuel, des sautes et un fort bougé de l'image noir/blanc (une sorte de *flash-back* qui s'interprète de façon intertextuelle en référence à d'autres films du cinéaste, où la TV et la radio étaient présentes), confère aux locuteurs ce statut de « morts-vivants », suggérant que les humains auront bientôt cessé d'« émettre ».

Tous les cas de marquage que j'ai discutés ici exhibent la nature mécanique de la parole enregistrée. Le régime dominant de la « transparence » consiste néanmoins à tout mettre en œuvre pour masquer l'intervention du pôle machinique du dispositif. À cet égard, la vocalité joue un rôle décisif : en humanisant la représentation, elle la rapproche de l'audiospectateur et fait oublier les conditions mécaniques dont elle résulte. L'exploitation du potentiel « personnalisant » des caractéristiques vocales s'effectue par la

739. L. Jullier, *Les Sons...*, *op. cit.*, p. 31.

740. M. Chion, *Un art sonore...*, *op. cit.*, p. 427.

simple individualisation de la voix-action de tous les personnages du film, mais aussi de deux façons particulières, chacune appartenant à l'un des pôles attraction/narration. Premièrement, elle peut s'appuyer sur un intertexte extracinématographique qui, lui, convoque de façon dominante la voix-attraction. C'est le cas de la référence aux spectacles «de type vaudeville» dont j'ai traité au chapitre précédent. Deuxièmement, la personnalisation vocale peut créer, par le biais de la voix-narration, ce que j'appellerai, en m'inspirant de la terminologie d'André Gaudreault, un «méga-énonciateur fictif». Afin d'évaluer comment cette «figure» anthropomorphe s'élabore dans l'esprit du spectateur, il me faut d'une part préciser le sens que j'assigne au qualificatif «fictif», d'autre part définir les diverses procédures de personnalisation sur lesquelles repose l'instauration de cette instance.

La prise en compte de la personnalisation d'une instance énonciative (verbale) interne au texte filmique (qu'elle soit diégétique, homodiégétique ou «extradiégétique») nous ramène au paradoxe théorique dont Jean Châteauvert a fait le point névralgique de son étude sur la voix-over. En effet, ce procédé est généralement utilisé pour donner l'impression que la voix provient de l'origine du discours filmique, ce qui a pour conséquence de conférer une unité à ce «foyer» et de faire comme si ce *produit* qu'est le film dépendait exclusivement d'une instance humaine. Alors que nous avons vu qu'il existe une grande diversité de niveaux de marquage, ce principe permet de nier la part machinique du dispositif. On l'observe tant dans la fiction que dans la non-fiction, même si la première a tendance à accentuer ces phénomènes afin de favoriser l'adhésion du spectateur à un monde qui présente une moins grande «complétude». En outre, le degré généralement supérieur de narrativité qui caractérise les films fictionnels a souvent pour corrélat une individualisation plus forte des protagonistes. Il n'est dès lors pas étonnant que Sarah Kozloff, qui étudie la voix-over dans le domaine restreint du film de fiction américain, soit particulièrement sensible à cette dimension rhétorique qui vise à donner l'impression que la voix-over commande l'ensemble des paramètres de la représentation. En dépit des précautions prises pour inscrire cette instance dans une hiérarchie et pour la détrôner sur le plan théorique du piédestal que la plupart des films semblent lui conférer – si bien que les écarts à cette norme comme ceux dont procède *La Lectrice* (Michel Deville, 1988) deviennent des faits marquants

pour un théoricien comme Châteauvert, Sarah Kozloff reconduit à mon sens une conception de la «voix narrative» qu'il s'agirait de prendre véritablement entre guillemets. Ainsi dresse-t-elle, à la fin de l'introduction théorique de son chapitre sur le «narrateur à la 1<sup>re</sup> personne», le bilan suivant :

«Pour résumer, un narrateur homodiégétique en voix-over est toujours subsumé par (et donc subordonné à) une instance narrative plus puissante, le grand imagier qui dramatise l'action sur l'écran. Mais en tant que spectateurs, nous ne désirons en général pas voir le grand imagier qui est moins définissable, moins familier [...] et nous percevons le personnage comme étant le principal narrateur.»<sup>741</sup>

Précisons d'abord que cette remarque a trait à un type particulier de narration verbale, où la voix est celle d'un narrateur homodiégétique. Or ce cas constitue, comme nous le verrons, le terrain d'exercice privilégié des procédures personnalisantes. Kozloff tient à minimiser la place de l'énonciateur en voix-over en postulant, comme je propose aussi de le faire, que cette position lui est octroyée par le spectateur. On pourrait motiver cette attitude spectatorielle en faisant l'hypothèse d'une quête de «confort», légitime dans le contexte d'un divertissement populaire comme le cinéma, nécessitant d'écarter l'«inquiétante étrangeté» résultant de la perceptibilité des manifestations machiniques du dispositif. Ce qui me paraît par contre discutable dans l'explication de Kozloff, c'est que la destitution théorique de la prééminence du narrateur *over* (établie empiriquement) passe par la proclamation d'une nouvelle instance faïtière qualifiée par le terme emprunté à Laffay, le «grand imagier», et non par une prise en compte de la diversité des matières de l'expression. Pourtant, Kozloff avait précédemment fait mention de celles-ci en disant que «la voix-over n'est seulement qu'un élément parmi d'autres, à l'instar de la musique, des effets sonores, du montage, de l'éclairage et ainsi de suite, à travers lesquels le texte cinématique est narré»<sup>742</sup>. Cependant, la théoricienne ne me semble pas tirer les conséquences de la multiplicité des lieux potentiels de marquage car, en englobant tous ces éléments

741. S. Kozloff, *op. cit.*, pp. 48-49.

742. *Ibid.*, pp. 43-44.

sous la responsabilité du « grand imagier », il n'est plus possible de les considérer dans leur spécificité. Par ailleurs, cette terminologie semble réinstaurer la préséance de la dimension visuelle, comme si la voix-over venait se greffer sur une image qui, par essence, serait première. En fait, au lieu d'écarter ou de relativiser le principe d'une instance unique, englobante et qui tend à l'anthropomorphisation, Kozloff, à l'instar de Châteauevert lorsqu'il recourt à la notion de « locuteur L », ne fait qu'établir une hiérarchisation qui évacue le problème du « site du film » en le déplaçant vers une instance extérieure aux manifestations textuelles sur lesquelles porte l'analyse. Selon moi, pointer de la sorte les instances énonciatives *fictionnelles* conduit à occulter le statut fondamentalement *fictif* qui est attaché à cette place centrale de l'énonciation, quel que soit le nom donné à l'instance qui l'occupe. Car ce n'est pas parce que le film utilise des êtres de fiction pour faire écran à la stratification complexe de sa structure énonciative que la théorie doit entreprendre la même éviction de la part machinique du dispositif cinéma.

En vue de parer à la confusion que l'on trouve couramment entre fiction et récit, j'avais proposé dans *La Fiction au cinéma* un distinguo entre la *fictivité* (et l'adjectif *fictif*), qui renvoie au caractère construit de la représentation (et à laquelle ressortit par exemple la mise en forme narrative), de la *fictionnalité* (et du caractère *fictionnel* de certaines données) qui concerne les propriétés sémantiques de la diégèse comparativement au monde d'expérience du spectateur<sup>743</sup>. La nécessité de cette distinction était bien sûr liée à l'objet de l'ouvrage, mais il me semble que cette opposition peut être d'un certain secours pour résoudre les interrogations soulevées ici quant à la place de la voix-over au sein de l'énonciation filmique. En effet, si les instances que le film désigne explicitement comme étant à sa source sont *facultativement* fictionnelles (ce sont des personnages individualisés grâce à des procédures spécifiques), la position qu'elles occupent implicitement dans l'organisation discursive du film est *nécessairement* fictive. Ainsi, la prise en compte d'un « méga-narrateur » proposée par certains théoriciens se situe au niveau de la fictivité. Or cette construction fictive convoque déjà, à un certain degré, un modèle anthropocentrique qui nous éloigne du marquage énonciatif des matières de l'expression

743. A. Boillat, *La Fiction au cinéma*, op. cit., pp. 31-33.

cinématographique, car celles-ci interagissent sans que s'en dégage un « foyer » unitaire. Par conséquent, cette origine fictive est assimilée à une instance fictionnelle (et même réelle lorsqu'il s'agit de l'« auteur » du film), pour peu qu'interviennent certaines procédures personnalisantes qui la parent d'attributs humains. Il faut certes que le spectateur accepte de jouer le jeu, mais cela ne pose généralement pas de problème puisque ce principe d'humanisation correspond aux aspirations profondes des consommateurs de productions audiovisuelles<sup>744</sup>. Sarah Kozloff souligne avec justesse le fait que la construction d'un narrateur homodiégétique répond à un désir sinon de fiction, du moins d'unité. À mon sens, la fictivité de tout « poste » d'énonciation permet non seulement d'expliquer pourquoi le spectateur d'un film en voix-over a l'impression d'assister à l'actualisation d'une sorte d'équivalent de la « voix narrative » du roman, mais aussi d'élargir l'effet « centripète » des différentes fonctions de la voix-over (homogénéisation, linéarisation, configuration, etc.) au-delà de la distinction fiction/documentaire. Dans l'un et l'autre cas, il y a création d'une « méga-énonciation fictive » qui, dans la fiction, peut s'incarner dans un narrateur fictionnel.

Cette conception du versant humain de la structure de l'énonciation filmique telle qu'elle est inférée par l'audiospectateur permet à mon sens de saisir la spécificité de la voix-over. À l'instar de toutes les autres voix présentes dans un film, elle contribue, sur la base de la matérialité vocale, à personnaliser le locuteur selon un processus qui, toutefois, s'avère moins influent que celui engagé par la voix synchrone, qui propose une « fusion » de la parole et du locuteur. Par contre, la voix-over dispose d'un potentiel énorme d'humanisation lorsqu'elle accède, à partir du plan de la méga-énonciation, à la création d'une instance qui, bien que fictionnelle, se trouve au sein d'une hiérarchie fictive dans une position de supériorité par rapport aux autres niveaux (ou instances créées à ces niveaux). Deux facteurs me paraissent décisifs dans l'assimilation de la voix-over au niveau fictif de la méga-énonciation : l'absence intrinsèque d'ancrage dans la diégèse visualisée et l'impression donnée que les images « naissent » des mots. C'est pourquoi la détermination du sens de la *relation de subordination* entre image et paroles *over* doit être envisagée sous l'angle de la

744. Voir le concept de « double magique » et l'analyse des fondements anthropologiques du désir d'une représentation « totale » de l'humain proposés par E. Morin, *Le Cinéma...*, *op. cit.*



construction d'une instance méga-énonciative. En faisant croire que les images se « plient » aux mots (notamment grâce à une certaine temporalisation de la coréférence et à la structure du récit rétrospectif), le film place l'énonciateur verbal dans une position fictive de domination. C'est pourquoi j'ai tenu à mettre en évidence les différentes manifestations de la voix en tant qu'instigatrice de mondes, que cela soit dans les prologues ou grâce à l'usage de temps verbaux passés.

Le passage du statut d'instance fictive à l'incarnation fictionnelle s'opère quant à lui *via* les procédés de la focalisation communément étudiés en narratologie filmique. Avec la mise en place d'un « méga-énonciateur », c'est l'ensemble de la représentation qui se donne comme le produit d'une intériorisation. La personnalisation et la subjectivisation gagnant la totalité du discours filmique, le spectateur se trouve face à un « visage » humain, à son double. Comme le spectateur se fantasme en tant que centre de la représentation, l'énonciation verbale *over* sur laquelle se fonde la domination de l'humain se fait l'équivalent du « discours intérieur ». La référence à la machinerie qui résulte d'éléments ponctuels de marquage de la matérialité des signifiants cinématographiques est ainsi masquée par l'identification à la voix. Lorsqu'il y a création d'un méga-narrateur fictionnel en voix-over dont le spectateur partage par ailleurs le « point de vue » (ce narrateur étant simultanément un personnage du film), on peut dire que le spectateur *se parle à travers le film*. L'effacement des conditions machiniques de production du film vaut dans ce cas également pour l'image, considérée comme l'émanation de l'expression verbale de cet énonciateur diégétique. Comme l'a relevé François Jost, le texte *over* « a une priorité effective dans la lecture narrative : la focalisation verbale *ancree* l'ocularisation »<sup>745</sup>. La monstration est par conséquent rapportée au regard d'un personnage, ce qui dissimule la présence de la caméra au profit de l'identification secondaire au personnage.

Notons que cette fictivité de l'instance d'énonciation verbale du cinéma parlant, que l'on doit prendre en compte indépendamment de la pluridiégéticité ou de la fictionalité de cette instance, ne s'applique pas au bonimenteur. Dans le cinéma parlé, la fictivité ne concerne que l'organisation de la bande-image, non le locuteur. En effet, celui-ci n'y est pas lui-même le produit d'un discours,

745. François Jost, *L'Œil-caméra*, Lyon : PUL, 1987, p. 82.

mais une présence effective soumise aux contraintes du dispositif. C'est pourquoi l'audiospectateur a le sentiment qu'*on lui parle du film*: la fonction de médiation du bonimenteur interdit que le spectateur s'identifie à ce dernier sur le mode qui caractérise la lecture d'un film à voix-over.

Dans un premier temps, on peut se concentrer sur les cas où l'identification de l'énonciateur verbal en voix-over par l'audiospectateur ne s'opère pas sur la base de déductions effectuées à partir d'informations diégétiques, mais grâce à l'activation d'un savoir d'ordre extrafilmique (voire extracinématographique). Nous avons déjà abordé deux types de relation verbo-iconique où les inférences intertextuelles se sont avérées primordiales: d'une part les voix-over qui, convoquant un certain degré de déliaison (par plaquage, dissociation ou doublage), incitent l'audiospectateur à associer la voix proférée à l'acteur visible à l'écran; d'autre part la voix pseudo-attractionnelle qui définit la position réceptrice en fonction des canons représentationnels issus des spectacles vivants. Comme je l'indique dans le schéma récapitulatif qui se trouve à la fin du présent chapitre, ces deux phénomènes donnent effectivement lieu à une personnalisation (l'intertexte est une sorte d'équivalent, dans le pôle de l'humain, du «savoir de l'arché» pour le pôle machinique), mais celle-ci n'est pas transférée fictivement sur le niveau méga-énonciatif. Occupant une position intermédiaire entre l'acteur réel et le personnage (comme Bogart au début de *La Comtesse aux pieds nus* ou Al Jolson dans *Le Chanteur de jazz*), le «sujet humain» n'est pas rapporté au contrôle de la fabrication même du film.

Il faut bien sûr distinguer la voix-attraction dont usaient bonimenteurs, «Monsieur Loyal» et chanteurs de la voix dont l'enregistrement même est exhibé en tant que divertissement. Je propose d'appeler «*son comme attraction*» ce pôle machinique. Les inférences qui s'y rapportent sont alors de l'ordre du marquage phonographique. Quant à la voix pseudo-attractionnelle, généralement diégétique et synchrone, sa constitution par détours intertextuels rejoint tout à fait ce que Marc Vernet dit du fonctionnement du regard-caméra dans la comédie musicale. Son explication permet de rejoindre les conclusions d'ordre énonciatif auxquelles j'ai abouti en analysant la voix dans les spectacles scéniques populaires:

«Mais je relève que dans le numéro de chant, c'est la vedette (dans notre exemple, Fred Astaire) qui prend le pas sur le personnage, c'est la star [...] Vedette de cinéma? Vedette de music-hall d'abord, parce qu'il est vrai que Fred Astaire en est issu et que le film le réintègre dans ce cadre [...]. Dans le regard à la caméra de type Astaire, j'accepte pour un temps de ne plus être au cinéma mais au music-hall où je sais que la vedette interpelle la salle. Du coup, ce n'est pas l'instance d'énonciation qui se dévoile et le film ne devient pas pour autant exhibitionniste. Au contraire, l'énonciation s'efface totalement au profit d'une autre, au profit d'un autre spectacle qui a ailleurs son origine: le numéro de music-hall. Si bien qu'au moment où le film semble changer de régime de convention, il n'en change absolument pas: il reste transparent.»<sup>746</sup>

On le voit, la façon dont Vernet conçoit le regard-caméra – équivalent dans ce cas de ce que j'ai proposé de nommer le «regard-public – correspond tout à fait à son «homologue» vocal, la voix-attraction qui, comme je l'ai montré, est associée par le spectateur à une pratique qui s'origine dans les spectacles scéniques «de type vaudeville». D'ailleurs, le regard-caméra constitue l'une des modalités de l'adresse au spectateur qui, fréquemment, se manifeste conjointement au régime vocal de la voix-attraction. Dans de tels cas, la star «perce» sous le personnage, ce qui confère à ce dernier une certaine humanisation et authentifie sa performance, mais lui interdit d'accéder à l'immatérialité qui définit une pure instance du discours (la «voix narrative» du film). Inversement – du moins eu égard à notre opposition entre l'humain et la machine, le décentrement qu'opère la référence à un spectacle non cinématographique désamorce, comme l'explique très bien Vernet, tout effet de réflexivité. En dépit du mode de l'adresse et de l'esthétique des attractions, cette voix ne renvoie pas, dans notre optique, à la matérialité du support filmique, et ne rompt pas non plus l'unité et la clôture de la diégèse. En fait, la voix-attraction révèle la capacité que possède tout locuteur à référer, par le biais d'une posture énonciative spécifique, à un autre contexte d'énonciation qui, dès lors, se voit implicitement convoqué avec l'ensemble de ses caractéristiques communicationnelles. Erving

746. Marc Vernet, *Figures de l'absence. De l'invisible au cinéma*, Paris: Éditions des Cahiers du cinéma, 1988, pp. 13-14 (je souligne).

Goffman appelle ce phénomène l'«enchâssement»: postulant que «toute énonciation et son audition portent la marque du cadre de participation au sein duquel elles ont lieu», le sociolinguiste précise que «ces marquages, nous pouvons les mimer et les rejouer», par exemple en insérant «dans un cadre de participation ce qui est structurellement marqué comme appartenant à un autre»<sup>747</sup>. Le terme de «marquage» que j'utilise ici pour désigner le renvoi à la matérialité du médium peut donc également être rapporté, au sens de Goffman, à un «cadre de participation» donné, c'est-à-dire à un ensemble de codifications régissant le positionnement des différents protagonistes de la communication. Dans ce cas, nous nous situons résolument du côté de l'humain. Dans le cinéma parlant, ces phénomènes sont néanmoins soumis à l'irréalisation de la *fictivité*, conséquence de ce que j'ai appelé l'«intégration de premier niveau». En fait, le marquage de l'énonciation filmique tel que je l'entends ici correspond bien plus à ce que Goffman qualifie de «bruits». Il ne s'agit toutefois pas de les concevoir nécessairement comme des perturbations (à moins qu'il soit question du «bruit du projecteur», qui ne résulte pas d'une intention de l'instance de production), puisque, constitutifs du médium, ils peuvent être eux-mêmes utilisés à des fins signifiantes. Goffman évoque d'ailleurs la composante technologique (parallèlement aux facteurs physiques et biologiques) en tenant compte de ce que j'ai proposé d'appeler, à la suite de J.-M. Schaeffer, le «savoir de l'arché», qui permet à l'audiospectateur de développer une conscience des niveaux auxquels les sons sont transformés: «Aujourd'hui, tous ceux qui regardent la télévision savent parfaitement bien que les problèmes de réception peuvent provenir de sources totalement différentes: du studio, de leur propre appareil, d'effets électroniques dans le voisinage [...].»<sup>748</sup> En fait, si la transparence de la matérialité phonographique est généralement assurée au cinéma, c'est parce que, «dans tout système de communication, les récepteurs acquièrent une tolérance à toute une série de bruits»<sup>749</sup>. La perception auditive quotidienne du langage parlé dans des environnements variés (salle bruyante où l'auditeur sélectionne une voix en particulier, «bruits» qui se sont accrus avec la téléphonie mobile) n'est pas sans habituer l'audiospectateur du cinéma à la

747. E. Goffman, *op. cit.*, p. 10.

748. J.-M. Schaeffer, *L'image précaire, op. cit.*, p. 189.

749. *Ibid.*, p. 190.

mise à l'écart d'éléments «perturbateurs» parce qu'étrangers à la voix humaine.

Deux autres origines «vocales» favorisent chez le spectateur des inférences qui, elles, ont une incidence sur la représentation de la méga-énonciation fictive dans la mesure où elles convoquent celui que le spectateur considère comme le responsable de ce qu'il voit: «l'auteur» (c'est-à-dire, à notre époque, le réalisateur du film). Nous revenons donc à la question traitée par Châteauvert (voir *supra*), mais non sans avoir bâti les fondements théoriques nécessaires à une relativisation de la place de cette instance dans l'ensemble des questionnements liés à l'énonciation filmique. La perception, dans la voix-over, de caractéristiques propres à un auteur appelle deux types d'intertexte: l'un est cinématographique (le cinéaste), l'autre littéraire (dans le cas de l'adaptation d'une œuvre d'un écrivain célèbre).

Bien plus que n'importe quel trait «stylistique» susceptible d'être rapporté au travail d'un réalisateur donné, sa voix représente, pour ceux qui la (re)connaissent, un gage de son inscription dans le discours filmique. Notons qu'indépendamment de l'identité du locuteur, le procédé de la voix-over est en lui-même à compter parmi les paramètres qui définissent le «style» d'un cinéaste lorsque celui-ci y recourt de façon systématique, qui plus est s'il en fait dans plusieurs films une utilisation similaire. Toutefois, lorsque cette voix lui appartient en propre, elle se fait porteuse de ce que Jean Châteauvert appelle un «effet de signature» beaucoup plus marquant<sup>750</sup>. Le réalisateur étant idéalement placé à la tête de l'ensemble des opérations de production de sens d'un film (qu'il ne faut pas confondre avec le fonctionnement effectif du discours filmique), le spectateur aura tendance à lui assigner la position faîtière de l'organisation énonciative fictive. Parce qu'orale, cette manifestation du cinéaste permet simultanément l'*auteurisation* et l'*autorisation* du texte: lui seul semble exercer son autorité sur la transmission du discours, ce qui est effectivement le cas pour le discours verbal, fictivement le cas pour le discours filmique dans son ensemble. Concrètement, le metteur en scène de cinéma est celui que la production désigne comme étant le *responsable* de la bonne marche de la réalisation du film (sauf bien sûr dans les cas mythiques de fiasco où le réalisateur d'un film hollywoodien, jugé «irresponsable», est remplacé en cours de tournage): il n'est donc pas étonnant de constater chez le spectateur ce

750. J. Châteauvert, *Des mots à l'image...*, *op. cit.*, p. 121.

réflexe anthropocentrique qui consiste à reporter cette donnée extrafilmique sur la question de la *responsabilité énonciative* interne au film, quand bien même cette dernière n'est qu'un effet de texte.

La «signature» du cinéaste (qu'il est nécessaire de concevoir, je le souligne, en termes d'«effet») s'impose néanmoins avec plus ou moins de force en fonction de deux critères: l'autodésignation et le marquage de la matérialité vocale. On distinguera un cinéaste comme Truffaut qui s'exprime *over* en tant que personnage (*L'Enfant sauvage*, 1970) ou en tant que figure non individualisée (le narrateur extradiégétique des *Deux Anglaises et le continent*) de Sacha Guitry qui, dans *Le Roman d'un tricheur*, se présente d'entrée de jeu (et de façon ostensible) comme le réalisateur du film. Dans ce second cas, il n'est nullement besoin que le spectateur active des inférences intertextuelles pour reconnaître l'origine de la voix; par contre, celles-ci lui permettent de mieux *connaître* cette voix, établissant ainsi un lien de proximité tout particulier qui s'apparente, sous une forme fixée, aux conditions de réception du cinéma parlé. La réflexivité filmique induite par l'autodésignation verbale du cinéaste n'a de ce fait pas pour effet de renvoyer à la matérialité du signifiant, mais au contraire d'humaniser l'origine de l'ensemble du discours filmique. Par contre, plusieurs théoriciens du cinéma ont émis l'hypothèse selon laquelle la présence de l'auteur à travers un ensemble de caractéristiques vocales marquées fonctionne comme un facteur de rupture de la transparence du récit filmique. Il est vrai que la voix s'affiche alors dans sa matérialité au lieu de s'effacer dans cette présence-absence qui constitue un trait spécifique de la «voix narrative». C'est précisément en ces termes qu'Anne-Sophie Nedelec discute la voix-over de Guitry et de Cocteau dans leurs propres films:

«Pour que chaque spectateur puisse la faire sienne, la voix narrative doit s'apparenter à un texte écrit qui serait entendu (parce que lu) dans le for intérieur de la lecture. Au contraire, si nous entendons qu'une voix-off s'écoute parler, l'image d'un corps, d'une personne vient empêcher l'identification à la voix: elle s'interpose entre l'image et nous au lieu de nous y introduire. Les voix-je de Cocteau et Guitry produisent un effet insolite et fascinant, voire envoûtant, car en même temps qu'elles portent la narration, elles l'encombrent de leur présence.»<sup>751</sup>

751. A.-S. Nedelec, *art. cit.*, p. 42.

J'abonde dans le sens de cette explication, sauf en ce qui concerne l'effet d'«interposition» car, selon moi, la voix contribue au contraire à nous introduire dans le monde du film sur le mode de la médiation. D'ailleurs, le qualificatif «envoûtant» suggère une certaine immersion, voire une adhésion émotive. Nedelec procède néanmoins à une comparaison intéressante entre la lecture d'un texte écrit auquel le lecteur prête intérieurement une certaine «voix» (c'est-à-dire un «discours intérieur» réalisé phoniquement tout en demeurant mental) et la perception d'une voix-over au cinéma. Le point commun entre la manifestation textuelle (virtuellement orale) d'un roman et la présence sonore de la voix-over réside dans la notion de «voix narrative». L'équilibre fragile de la présence-absence est rompu par la vocalité, indice d'une présence indubitable, physique. Nedelec est d'avis que la neutralité des caractéristiques vocales est nécessaire afin que tout le monde puisse se «projeter» dans la voix-over, se l'approprier comme étant la sienne. Toutefois, cette appropriation rencontre une limite qui réside dans l'indépassable différence existant entre le texte écrit et la voix-over: ce «qu'entend» le lecteur en son for intérieur, c'est toujours, ainsi que l'a souligné François Jost (à propos de l'usage de la troisième personne, mais cela vaut à mon sens pour tout texte littéraire), *sa propre voix*<sup>752</sup>, alors que le spectateur d'un film est confronté à la voix de l'Autre, de celui qui lui montre les images ou de celui qui appartient au monde du film.

Lorsque la voix-over est celle d'un personnage, l'audiospectateur se soumet à une double identification: il s'identifie à la voix pour s'identifier au personnage, et inversement. Ce processus s'opère dans une dynamique de réciprocité que déterminent les différents degrés de déliaison de la voix-over. Il n'en reste pas moins qu'au cinéma, la matérialité vocale peut s'afficher de façon plus ou moins ostentatoire. Sur ce point, Châteauvert partage l'avis de Nedelec lorsqu'il affirme que «le discours marqué par le grain ou le timbre de la voix devient ainsi un discours dont la valeur pragmatique est compromise, un discours qui a pour premier effet de faire entendre sa matérialité<sup>753</sup>. Chez Châteauvert, cette explication conclut une section intitulée «La voix marquée», à laquelle fait suite une brève partie consacrée à «La voix de l'auteur». Le lien

752. F. Jost, *Un monde à notre image, op. cit.*, p. 97.

753. J. Châteauvert, *Des mots...*, *op. cit.*, p. 121.

entre l'une à l'autre n'est cependant pas explicité. Si, dans le cas de la voix d'un personnage, la matérialité vocale peut à la rigueur établir une séparation entre le spectateur et le film, la voix auctoriale renforce au contraire sa « valeur pragmatique » en instaurant avec le spectateur une relation qui, certes, diffère du pacte fictionnel traditionnel, mais relève néanmoins d'une démarche contractuelle. La reconnaissance de la voix de l'auteur par le spectateur s'effectue grâce à des inférences péritextuelles (la mention du cinéaste au générique ou son autoprésentation dans le générique) et intertextuelles (cette voix a déjà été entendue à la radio, dans des entretiens télévisuels, dans d'autres films du même cinéaste, dans des citations parodiques, etc.) qui induisent l'attribution de l'énonciation filmique à cette figure humaine. Par conséquent, il me semble plus juste de parler d'identification *de* la voix plutôt que d'identification « à » la voix. Certes, un « corps » s'interpose dans la représentation mentale du spectateur, mais il est associé à une force agissante, à un élément central de la genèse du film qui, nécessairement, « fait corps » avec ce dernier. On sait que ce corps n'est pas là pour se donner à voir (même s'il se donne à entendre) dans les moments en voix-over, mais pour donner à voir.

Il faut en outre distinguer des cinéastes qui, comme Guitry dans le récit-cadre du *Roman d'un tricheur*, apparaissent de surcroît dans le monde du film en tant que sujet parlant et agissant. Ce cas de figure combine deux paramètres : le locuteur *over* est à la fois l'auteur *et* un personnage (même si ce personnage incarne à son tour l'auteur). De tels films jouent ainsi sur deux tableaux, créant avec le public un rapport d'intimité particulier : à l'identification au personnage s'ajoute l'identification de l'auteur, ce qui permet de concilier la cohérence diégétique avec l'authentification de la référence extrafilmique. Qu'en est-il de l'effet d'« interposition » évoqué par Nedelec lorsque la voix n'est pas celle du cinéaste ? Je ne pense pas que l'exhibition de la matérialité vocale, fondamentalement différente à cet égard de la matérialité phonographique, tende systématiquement à enrayer l'identification au personnage. Il me semble plutôt que l'inverse se produit : les caractéristiques vocales étant propres à un individu, elles contribuent à personnaliser l'instance d'énonciation verbale. Dans *L'Année dernière à Marienbad* (Alain Resnais, 1961), les personnages interprétés par des acteurs qui demeurent la plupart du temps totalement impassibles se construisent principalement par opposition différentielle



des timbres, des accents, des débits et des intonations: en dépit d'un faible degré de psychologisation (rappelons que dans son scénario, Robbe-Grillet se contentait de se référer aux protagonistes grâce aux lettres A, M et X), chaque personnage est individualisé, «coloré» par l'expression vocale. C'est pourquoi la vocalité, qu'elle soit imputée à une instance auctoriale ou à un personnage, est à compter, du point de vue de la structure de l'énonciation filmique, parmi les procédures de personnalisation du locuteur qui renforcent la construction fictive du «pôle humain».

Le spectateur peut être amené à créer sur une base intertextuelle une autre figure d'«auteur» dans le cas particulier où le film est une adaptation d'un texte littéraire (ou comprend une voix-over qui cite un tel texte). La reconnaissance ne porte dès lors plus sur les caractéristiques vocales, mais sur le texte *over* proprement dit. En effet, ce dernier peut reprendre «à la lettre» le texte connu du spectateur tout en l'actualisant dans l'oralité, ce qui, pour certains spectateurs, est déjà une intolérable personnalisation par rapport à la voix qu'ils avaient attribuée en tant que lecteurs, dans leur «for intérieur», à tel ou tel personnage. Il faut remarquer que, dans les films de fiction, l'adaptation d'œuvres romanesques est probablement le lieu d'application le plus fréquent du procédé de la voix-over. Sarah Kozloff consacre significativement une section de son ouvrage à la contestation du lieu commun voulant que la voix-over soit un procédé foncièrement littéraire. La remise en cause qu'elle propose est fondée, mais il s'agit de prendre également en compte ce stéréotype en tant que fait de réception: spontanément, de nombreux spectateurs associent la présence de la voix-over à l'existence d'une source littéraire. On comprend donc que l'activation de cet intertexte soit une donnée importante de la réception des films en voix-over. Au niveau textuel, on peut en outre s'interroger sur la nature du «déguisement» du narrateur filmique *over* en narrateur de roman dont parle Kozloff<sup>754</sup>. La théoricienne revient sur ce point lorsqu'elle étudie la narration à la 1<sup>re</sup> personne et mentionne, parmi les différentes fonctions de celle-ci, la capacité à «recréer la voix narrative du roman ou [à] s'y référer»<sup>755</sup>. Si on rapporte cette activité de «référence» au texte proprement dit, on comprend que celui-ci, une fois intégré au film, continue d'apparaître à

754. S. Kozloff, *op. cit.*, p. 19.

755. *Ibid.*, p. 41.

l'audiospectateur averti comme étant la « propriété » de l'écrivain. La médiation du réalisateur ou du scénariste du film est bien sûr inévitable, ne serait-ce que dans le choix de laisser le texte original inchangé ou de limiter les modifications qu'il subit. Ainsi le cinéaste François Truffaut, dont on a déjà relevé à plusieurs reprises l'intérêt pour la dimension vocale, écrit-il à propos de la genèse de son film *Jules et Jim* (1962), dans lequel la voix-over occupe une place essentielle :

« Dès les premières lignes, j'eus le coup de foudre pour la prose d'Henri-Pierre Roché [auteur du roman homonyme]. [...] Pendant le tournage et le montage du film, il m'arrivait fréquemment de repousser le scénario, de rouvrir mon exemplaire du roman et de noter telle ou telle phrase splendide à « sauver absolument », c'est-à-dire à intégrer dans la bande-son du film. »<sup>756</sup>

De tels cas d'« intégration » directe de phrases tirées d'un roman publié avant la sortie du film (l'ouvrage de Roché paraît treize ans avant cette adaptation) – et cela d'autant plus pour des « classiques » de la littérature que les spectateurs sont susceptibles d'avoir lus à l'école (à l'instar de *Madame Bovary*, notamment dans son adaptation par Chabrol en 1991) – ou lorsqu'il s'agit d'écrivains déjà passablement reconnus et identifiables stylistiquement comme Marguerite Duras pour *Hiroshima mon amour*, l'inférence intertextuelle contribue nécessairement à humaniser l'instance d'énonciation construite par le spectateur. L'effet de cette opération ne consiste pas forcément à favoriser l'immersion du spectateur dans le monde du film : la « littérarité » est généralement associée à un moyen de mettre à mal l'illusion référentielle. Mais quels que soient les a priori du spectateur quant à l'utilisation littérale d'œuvres scripturales, il n'en rapporte pas moins la source du texte proféré à l'écrivain dont il a reconnu la production. La figure d'énonciateur élaborée de façon intertextuelle est alors bifide (et c'est probablement cette scission qui fait obstacle à l'adhésion totale à l'univers filmique) : « l'auteur » est à la fois l'écrivain (pour le texte *over*) et le cinéaste (pour le reste), partition qui s'annule bien sûr dans les films réalisés par Robbe-Grillet ou Duras.

756. F. Truffaut cité par J. K. Sanaker, *art. cit.*, p. 99.

Ces auteurs qui «font autorité» grâce à leur emprise plus ou moins directe sur une partie du discours filmique demeurent des figures d'altérité que le spectateur, se sachant appartenir au même monde, reconnaît sans s'identifier à elles. Dans la fiction, la personnalisation est donc plus forte lorsque le spectateur est invité à occuper imaginativement la «place» d'une instance fictionnelle (fictivement) présentée comme le «foyer» humain de l'ensemble de l'énonciation filmique. À la question de la personnalisation s'ajoute alors celle de la définition d'un «point de vue», de l'expression d'une subjectivité. Ces questions ayant déjà été passablement traitées dans le champ de la narratologie du cinéma (voir notamment Branigan et Jost), je me limiterai à quelques aspects qui concernent en propre les narrateurs *over*, d'autant plus que les études consacrées à la dimension émotive du cinéma ou à la question de l'expression de la «subjectivité» ont tendance à écarter le procédé de la voix-over, peut-être parce qu'il est jugé trop peu «spécifique» au médium ou trop évident dans ses effets. Pour prendre l'exemple d'une approche d'inspiration cognitiviste, on constate qu'Edward Branigan signale d'entrée de jeu, dans son chapitre intitulé «*Subjectivity*», qu'il ne prendra pas en compte des «cas hybrides» comme ceux qui présentent un narrateur en voix-over appartenant à une situation-cadre<sup>757</sup>. En ce qui concerne la façon dont le narrateur-over médiatise sur les plans émotionnel, idéologique, axiologique, etc. le rapport du spectateur au film, un vaste champ théorique reste à défricher<sup>758</sup>. Je me contenterai d'établir les paramètres qui me semblent intervenir dans le processus d'anthropomorphisation d'une instance considérée par le spectateur comme relevant du niveau méga-énonciatif.

Mon but n'est pas de proposer ici une nième typologie des voix-over au cinéma, mais de concevoir la diversité des mises en œuvre auxquelles ce procédé peut donner lieu à travers deux axes principaux qui, selon moi, font de l'intégration de la parole orale dans un film un véhicule privilégié de l'humanisation de l'énonciation filmique. Il s'agit, d'une part, de la *personnalisation* résultant de

757. Edward Branigan, *Point of View in the Cinema*, Berlin ; New York ; Amsterdam : Mouton, 1984, p. 73.

758. Il est en effet étonnant de constater que les index thématiques de deux ouvrages des années 90 consacrés à la dimension émotionnelle de la réception filmique dans une optique narrato-cognitiviste, *Engaging Characters* de Murray Smith (New York : Oxford University Press, 1995) et *Emotion and the Structure of Narrative Film* d'Ed. S. Tan (Mahwah : Lawrence Erlbaum, 1996), font la part belle à la perception visuelle, mais ne mentionnent pas le terme «voix».

l'identification du locuteur, d'autre part, de la *subjectivisation* qui, s'élaborant généralement sur la base de la première, permet d'associer cette identité fictionnelle à une conscience, à un filtre par lequel semble nous parvenir la représentation de l'univers diégétique.

Qu'il soit visible ou non, le narrateur se construit progressivement en tant que personnage dans l'esprit du spectateur à travers la façon dont il s'exprime verbalement et grâce aux particularités de l'audiovisualisation du monde qui semble résulter de sa profération. Dans la conception de l'énonciation filmique prônée entre autres par André Gaudreault, le personnage-narrateur est considéré comme une instance déléguée dont l'activité narrante s'effectue à un niveau second. Si par contre nous abandonnons l'idée qu'il existe un «niveau premier» (méga-énonciatif) en dehors même de l'attribution d'une fonction de narrateur à un tel personnage, le «méga-narrateur» se réduit à une place vacante que tout «sujet parlant» peut fictivement occuper, notamment sous la forme d'êtres fictionnels. En octroyant fréquemment une position de savoir et de maîtrise au narrateur verbal sous la forme d'un récit rétrospectif, les films de fiction favorisent l'assimilation de la méga-énonciation fictive au narrateur fictionnel. Dans ma discussion de la notion d'extradiégéticité, j'ai montré que le narrateur n'est jamais extérieur à la diégèse, fût-il réduit à un énonciateur verbal peu présent. L'«extradiégéticité» ne peut résulter que d'une comparaison entre deux niveaux de narration qui forment une diégèse identique, ou une mégadiégèse lorsque le film procède à un emboîtement de mondes régis par des lois différentes. Jean Châteauevert a souligné ce point en affirmant que le narrateur *over*, quel que soit son statut, «apparaît comme un personnage toujours déjà inclus dans la fiction»<sup>759</sup>. Dans les films fictionnels, les narrateurs extradiégétiques sont donc nécessairement des «personnages», des entités inscrites dans l'univers diégétique. C'est pourquoi la matérialité vocale et certaines modalisations du discours verbal suffisent à assigner à l'instance *over* la consistance d'un être de fiction. Comme on l'a vu, ce n'est pas sur le couple intradiégétique *vs.* extradiégétique (et métadiégétique) qu'il s'agit de mettre l'accent, mais sur l'opposition homodiégétique *vs.* hétérodiégétique, c'est-à-dire sur la question de la présence du *personnage-narrateur* en tant que *personnage agissant* dans le récit narré.

759. J. Châteauevert, *Des mots...*, *op. cit.*, p. 44.

On comprend aisément que la personnalisation bénéficie de ce dédoublement puisqu'il permet d'augmenter le nombre d'attributs fictionnels implicitement conférés à l'instance fictive. Or, comme le note Genette, «l'absence est absolue, mais la présence a ses degrés»<sup>760</sup>. C'est pourquoi il me paraît possible de corréler les échelles de personnalisation et de subjectivisation avec le *degré d'homodiégéticité du narrateur*. Au cinéma, cette conception graduelle convoque à mon sens deux facteurs souvent interdépendants: la fréquence d'apparition du narrateur en tant que personnage visualisé dans son propre récit (ou, sur un mode mineur, la fréquence de la mention de celui-ci dans les paroles échangées par d'autres personnages, comme dans *Chaînes conjugales*) et la centralité de la place qu'il y occupe. Cette dernière peut être estimée selon une hiérarchie dont le plus haut point réside dans la fonction d'actant-sujet (principal et constant) correspondant au cas du «récit autodiégétique» de Genette. Si l'on ajoute à ces deux paramètres qui ont trait au récit enchâssé la fréquence de la manifestation *over* du narrateur, et en particulier le nombre d'énoncés caractérisés par la présence d'un «je» grammatical, on obtient un instrument permettant l'évaluation «quantitative» du degré de présence de cette instance. L'effet combiné de ces critères fonctionne d'autant mieux que le spectateur a tendance à rapporter les informations provenant des deux niveaux à une entité humaine unique.

Cette concentration des données sur un même personnage n'est cependant possible qu'en raison d'une fonction bien particulière de la matérialité vocale que je propose de nommer «*attestation d'homodiégéticité*». En effet, lorsqu'il s'agit de la voix-over d'un personnage du film, la comparaison permettant au spectateur d'identifier l'occurrence vocale n'est pas uniquement d'ordre intertextuel comme dans le cas des voix «auctoriales» envisagées ci-dessus, mais repose également sur des données strictement filmiques. François Jost a relevé à cet égard un aspect qui me paraît revêtir une grande importance du point de vue de la personnalisation de l'instance d'énonciation verbale et de la comparaison entre texte écrit et voix-over. Alors que le degré d'homodiégéticité (ou respectivement d'hétérodiégéticité) dépend en priorité dans un roman de l'utilisation de la personne grammaticale («je» *vs.* «il»), le cinéma

760. G. Genette, *Figures III*, *op. cit.*, p. 253.

fait intervenir la matérialité de la voix, puisque «c'est la reconnaissance du timbre de la voix du comédien-narrateur dans celle d'un acteur-personnage qui permettra de l'ancrer [dans la diégèse filmique] et non l'emploi du *je* ou du *il*»<sup>761</sup>. Bien que l'usage d'un pronom personnel donné ne me paraisse pas sans incidences sur le degré de personnalisation attribué au niveau *over*, la remarque de Jost est tout à fait pertinente en ce qui concerne la relation entre la voix-over et la représentation audiovisuelle. La personnalisation de l'énonciation narrative dépend en effet de l'association éventuelle de la voix-over avec un personnage visualisé. Il est intéressant de noter que François Jost a appliqué dans un autre texte ce même critère au bonimenteur en proposant de confronter sa voix à l'aspect physique du locuteur de l'écran<sup>762</sup>. On a vu toutefois que la fonction de médiation du bonimenteur n'a pas pour objet d'introduire le spectateur au monde du film sur le mode de l'identification au personnage, mais en facilitant son intégration à un dispositif machinique par l'introduction d'une composante propre aux spectacles vivants. La personnalisation ne passe donc pas par l'ocultation de «l'être-là» du bonimenteur, mais au contraire par l'effet de présence induit par l'exercice de la performance vocale.

L'homodiégéticité est une notion graduelle qui, à un niveau qualitatif, est également déterminée par le degré de subjectivité afférant à la représentation filmique du narrateur-personnage. Selon Branigan, la subjectivité au cinéma peut être définie comme «une instance ou un niveau spécifique de narration où la transmission est *attribuée* à un personnage du récit et reçue par nous *comme si* nous étions dans la situation du personnage»<sup>763</sup>. Cependant, «être dans la situation» du personnage constitue une donnée qu'il paraît souvent difficile de prendre en compte en dehors du célèbre cas du «plan subjectif» («*POV-shot*») auquel Branigan accorde une grande attention, et cela d'autant plus pour un narrateur *over* dont la «situation» ne peut être saisie en des termes spatiaux. Dans un article sur le film noir, Christine Brinckmann s'interroge sur la fonction subjectivisante de la voix-over, un procédé dont l'essor est lié selon elle à l'apparition, dans les années 40, d'un intérêt croissant pour la psychologie humaine et les expériences intimes de

761. F. Jost, *Un monde à notre image*, op. cit., p. 99.

762. François Jost, «Les mots pour le voir», in Francesco Pitassio et Leonardo Quaresima (dir.), *Scrittura e immagine*, Udine: Università degli Studi di Udine, 1998, p. 31.

763. E. Branigan, op. cit., p. 73.

l'individu. Elle constate dans le corpus qu'elle étudie un paradoxe qu'il ne s'agit pas de rapporter à tous les types de productions en voix-over (certaines offrant de surcroît un fort degré de modalisation du discours qui augmente de façon ostensible le degré de subjectivité), mais qui peut nous donner un indice du rôle joué par la voix-over relativement à la subjectivisation de la représentation : bien que ces films confèrent au spectateur une forte impression de subjectivité, l'examen révèle, selon Brinckmann, qu'ils sont dominés, en termes de pratiques filmiques, par une « neutralité » qui touche autant le style visuel que les paroles proférées *over*. L'auteur en déduit que la charge subjective résulte de la présence même de la voix-over à la première personne, et cela d'autant plus lorsqu'elle apparaît dès les premières images et médiatise notre entrée dans le film<sup>764</sup>. S'immerger dans la fiction, c'est, dans de tels cas, pénétrer les pensées d'un personnage. Il arrive même parfois que la voix-over retentisse avant la monstration de tout élément diégétique, émergeant du noir initial pour trouver en ce lieu de seuil filmique, ainsi que le dit Jean-Louis Leutrat, « un point commun, une connivence avec les spectateurs du film qui, eux aussi, sont dans l'obscurité »<sup>765</sup>.

Dans la fiction, où nous accédons aisément à l'intériorité des personnages (ainsi que l'a montré Käte Hamburger), la voix-over représente un moyen privilégié pour intérioriser ce qui est montré. Elle permet au médium d'échapper à la contrainte imposée par le fait que, pour reprendre les termes de François Jost, « le paraître est le régime ordinaire du cinéma »<sup>766</sup>. On peut donc dire que, dans un film fictionnel, le recours à la voix-over est, en tant que tel, un phénomène *marqué*. La perméabilité fantasmatique des consciences – ou du moins d'une seule, unique, « focale » – est exacerbée par l'immédiateté de l'oralité de la voix-over qui supprime l'altérité (du personnage et du narrateur) par l'actualisation, dans l'oralité et par le truchement du langage articulé, du « discours intérieur » du spectateur. Nombreux sont les théoriciens qui ont traité de la voix-over en tant qu'instrument introspectif. Par exemple, Béla Balázs considère à la fin des années 40 (période qui, comme on l'a vu chez Laffay, Chartier et Brinckmann, a connu une

764. C. N. Brinckmann, *art. cit.*, p. 121.

765. Jean-Louis Leutrat, *Vie des fantômes. Le fantastique au cinéma*, Paris : Éditions des Cahiers du cinéma, 1995, p. 103.

766. F. Jost, *L'Œil-caméra, op. cit.*, p. 73.

vogue de films en voix-over) que le son «asynchrone» (c'est-à-dire non synchrone, mais synchronisé) constitue pour le cinéma «la voie de son évolution future». Il s'explique ainsi :

«Dans le cinéma parlant asynchrone, l'action peut progresser sur deux plans parallèles: celui du son et celui de l'image. Par exemple, je vois le déroulement extérieur des événements, et entre-temps j'entends ce que les hommes pensent et sentent dans leur for intérieur. Ou inversement: le récitant nous dit les choses qui arrivent. Les voir serait superflu. Ce que je vois, c'est la série des *associations intérieures* – le processus irrationnel des associations d'images.»<sup>767</sup>

Dire ce qui n'est pas montré (parce que superflu) ou ne peut l'être (parce qu'intériorisé): telle semble être la vocation de la voix-over. La première option est majoritairement exploitée dans le documentaire, la seconde dans la fiction. Balazs met notamment l'accent sur le dépassement de l'extériorité qu'induit l'autonomisation du son par rapport à l'image. Il ne conçoit néanmoins pas l'intériorisation comme un phénomène interne à la représentation, mais la rapporte à la relation instaurée entre le spectateur et le film. Cette perspective le conduit à considérer le montage comme le déclencheur d'«associations intérieures». On retrouve l'idée de la voix du film comme verbalisation du «discours intérieur» d'Eikhenbaum: l'intériorité n'est pas uniquement celle du «for intérieur» des personnages de l'écran, mais aussi celle du spectateur qui déchiffre le langage des images. La voix-over se fait alors le point de convergence de deux intériorités, l'une fictionnelle (celle du personnage), l'autre factuelle (celle du spectateur). C'est pourquoi elle est le parfait vecteur de l'identification du spectateur au personnage.

Dans une comparaison entre récit littéraire au récit filmique, François Jost situe le cinéma du côté de la perception, le roman du côté de la pensée. Selon lui, l'avantage du second réside dans le fait que «l'expression de la *pensée* d'un personnage est beaucoup plus simple dans la mesure où il s'agit, avant tout, *de monnayer du verbal dans du verbal* (si l'on met de côté le cas des mouvements psychiques non verbalisés que l'auteur doit traduire ou transposer)»<sup>768</sup>.

767. Béla Balazs, *Le Cinéma*, Paris: Payot 1979 [1949], p. 209.

768. F. Jost, *L'Œil-caméra*, *op. cit.*, p. 17 (je souligne).



Afin que l'organisation visuelle du récit filmique puisse exprimer la pensée, il faut donc qu'elle s'adjoigne l'apport d'un véhicule sémiotique isomorphe à celui du récit littéraire: le langage verbal. La voix-over, comme le boniment ou les intertitres, permet donc de vaincre la contrainte de l'extériorité. Significativement, la plupart des cas de « focalisation interne » mentionnés par Jost (et notamment l'exemple dans *M. Arkadin* de Welles) font intervenir de façon prédominante la voix-over. En outre, ce procédé est constamment convoqué lorsque Jost conçoit une forme « idéale » d'expression de l'intériorité dans un film<sup>769</sup>. Le statut *over* provoque une dissociation de l'image du locuteur et de la voix qui permet d'isoler le plan psychique du plan perceptif, et donc de filtrer de façon optimale les informants transmis au spectateur. Il n'est donc pas étonnant que la « focalisation » s'en trouve affectée. En se mettant au service de la personnalisation, la voix-over participe à l'entreprise de masquage de l'origine mécanique de la représentation. À la subjectivisation de la représentation s'ajoutent, de façon souvent plus explicite et ponctuelle, des signaux qui se donnent à interpréter comme l'expression d'une subjectivité. On rejoint alors l'analyse des linguistes: l'ensemble de la représentation est placée *de facto* sous le sceau de la subjectivité parce que le locuteur établit une « attitude propositionnelle » par rapport à son discours (*/croire que/, /penser que/, /rêver que/, etc.*) et introduit dans celui-ci des modalisateurs. C'est même l'une des thèses les plus intéressantes de l'ouvrage de Châteauevert que de considérer l'influence de la voix-over sur l'image comme relevant non pas (prioritairement) d'un rapport temporel, mais d'une modalisation du « lacis audiovisuel »<sup>770</sup>.

Après l'indispensable détour par la machine, nous aboutissons ainsi au « sujet embrayeur » cher à Châteauevert et, avant lui, à Michel Colin (qui emprunte cette notion à Jakobson). Colin conçoit en effet cette question en relation avec le processus metzien de « l'identification secondaire au personnage »: l'« embrayeur » permet de « désigner un élément qui permet la mise en marche de la < machine fictionnelle »; « le spectateur [...] se reconnaît à travers ce personnage [...] comme faisant partie de la communication sociale instaurée par le film »<sup>771</sup>. On retrouve par conséquent la fonction de socialisation

769. *Ibid.*, par exemple p. 80 et p. 128.

770. J. Châteauevert, *Des mots...*, *op. cit.*, pp. 68-72 et p. 174.

771. M. Colin, *op. cit.*, p. 39.

qu'investit l'oralité au cinéma lorsqu'elle n'est pas assignée à une source totalement assujettie à la diégèse visualisée: le « discours intérieur » est rendu audible, manifeste, et en cela abolit la frontière entre l'intérieur du sujet (sujet percevant qu'est l'audiospectateur, sujet parlant qu'est le narrateur du film) et l'extérieur (le monde audiovisuelisé de l'enchâssement, mais aussi, pour le spectateur, le monde du film qui inclut ce narrateur). On comprend dès lors qu'aux yeux et aux oreilles de l'audiospectateur, l'accès au film s'en trouve totalement humanisé.

Ce que l'on a coutume d'appeler le « monologue intérieur » du personnage paraît constituer le degré le plus élevé de subjectivisation et instaurer le rapport le plus immédiat au spectateur. Comme la plupart des théoriciens, Jean Châteaupert distingue ce type de manifestation vocale de la « narration en voix over » par le fait qu'il « se conjugue généralement *au présent* et révèle les réflexions, angoisses, doutes, etc. qu'éprouve un personnage *dans une situation* »<sup>772</sup>. Cette définition fonctionne si l'on considère ce qui se produit « généralement », mais il ne me semble pas possible de statuer sur un ensemble fixe de traits appropriés à une description de cet emploi spécifique de la voix-over, notamment en ce qui concerne l'usage des temps verbaux. En effet, il y a des narrations en voix-over différentes du monologue dont le texte est conjugué au présent, et rien n'exclut a priori qu'un monologue intérieur ne soit raconté au passé (le film *Detour* de E. G. Ulmer se situe bien souvent à la frontière entre les deux régimes). Châteaupert a d'ailleurs tout à fait conscience de la perméabilité de ces catégories puisqu'il introduit plus loin, confronté à l'étude du *Plaisir* (Max Ophuls, 1951), la notion de « monologue intérieur narrativisé »<sup>773</sup>. Dans les cas de « plaquage » (voir VI.4), la voix est par contre plus proche du monde visualisé et donc plus fortement soumise aux lois diégétiques que les autres types de voix-over, de sorte qu'elle ne bénéficie pas de la distance et de la maîtrise (apparentes) dont dispose fictivement le méga-énonciateur. En dépit d'une forte subjectivisation, le monologue intérieur ne contribue donc pas à personnaliser l'instance fictive de l'énonciation filmique, mais, à l'instar de la voix synchrone, n'individualise un personnage qu'en étroite relation avec la diégèse.

772. J. Châteaupert, *Des mots...*, *op. cit.*, p. 143.

773. *Ibid.*, p. 168.

On peut illustrer les enjeux de cette différence avec le récit rétrospectif de Jeff (Robert Mitchum) dans *La Griffre du passé* (*Out of the Past*, Jacques Tourneur, 1947). La partie consacrée à la rencontre de Kathie à Acapulco (la femme qu'il est chargé de retrouver pour le compte de son riche fiancé, et dont il s'éprend) présente en effet de nombreux verbes décrivant des processus intérieurs qui apparentent cette voix-over à un monologue intérieur («narrativisé» puisqu'il est conjugué au passé). La voix-over de ce récit enchâssé au «je» à fort degré d'homodiégéticité (Jeff est l'actant-sujet, et notre savoir sur le monde du film est réduit au sien) rend tout de suite compte de la séduction exercée par Kathie sur le locuteur en suggérant l'état émotionnel qui fut le sien à ce moment-là («je la vis... et je compris pourquoi...»). Par exemple, le narrateur précise *over* qu'il s'était surpris lui-même à se réjouir de la fermeture de l'office télégraphique qui lui interdisait de faire savoir à son employeur qu'il avait retrouvé sa fiancée: «J'en fus heureux... et je compris soudain pourquoi.» Un complément de temps comme «soudain» procède d'une stratégie qui consiste à mettre le spectateur à la place du «sujet pensant» dans le «présent» des événements représentés visuellement. Néanmoins, le temps verbal rappelle à tout instant que le narrateur en sait plus que le personnage montré à l'écran. Cette question du surplus de savoir transparait littéralement dans le texte *over* tout en créant une ambiguïté entre le présent et le passé, par exemple dans les phrases suivantes:

- J'allai chez Pablo cette nuit-là. Je savais que j'y irais toutes les nuits jusqu'à ce qu'elle s'y montre. Et je savais qu'elle le savait.
- Je savais qu'elle ne viendrait pas le premier soir.
- Mais le soir suivant, je savais qu'elle viendrait.

La clairvoyance et la prescience du personnage du récit enchâssé ne peuvent être considérées indépendamment du savoir du narrateur qui, lui, connaît l'issue du récit: s'il peut dire qu'à l'époque *il savait*, c'est parce qu'aujourd'hui, *il sait*. La différence est donc notable entre cette situation et un monologue intérieur qui énoncerait, sur une image du détective attendant au bar, «je sais qu'elle viendra demain», car le pouvoir assertif de cette phrase demeure totalement dépendant des contingences diégétiques, le spectateur pouvant tout aussi bien penser que le personnage se

fait des illusions. Ici, par contre, l'énonciation verbale acquiert une dimension performative par rapport au cours ultérieur du récit: le locuteur affirme son emprise sur l'audiovisualisation. En effet, le passage de la deuxième phrase citée ci-dessus à la troisième est marqué au niveau de l'image par des mouvements d'appareil ostensibles et un fondu qui signifient une ellipse temporelle. L'illusion est de ce fait donnée que cette emprise du verbal s'étend sur la monstration et le montage. Le film de Tourneur, comme beaucoup d'autres, construit ainsi une position de maîtrise discursive, interprétée par le spectateur comme la trace d'une instance fictive d'énonciation que les marques d'intériorisation du texte permettent d'humaniser, tandis qu'elles occultent l'énonciation matérielle du discours filmique.

Dans le présent chapitre, j'ai tenté de développer une approche théorique de l'énonciation filmique informée par une réflexion sur la nature des voix du parlant comparativement à la parole vive du cinéma parlé. Dans le schéma récapitulatif que je propose ci-dessous, on retrouve donc, entre les deux pôles de la machine et de l'humain, les différents régimes vocaux étudiés qui, en fonction des paramètres traités, renvoient par inférences spectatorielles à l'un de ces pôles. Afin d'indiquer le rôle spécifique de la voix sur lequel ce schéma met l'accent, je l'ai distinguée du texte qu'elle véhicule et considérée dans ses deux dimensions matérielles (phonographique et vocale).

L'interaction entre la voix et l'image apparaît également dans ce schéma, puisque certains principes de synchronisation renvoient soit à l'humain (l'unité créée par la voix synchrone), soit à la machine (certains cas de déliaison vocale). La synchronisation s'effectue en outre sur l'axe vertical qui relie les différentes matières de l'expression sonore. Les deux grands faisceaux de déterminations qui influent sur les inférences du spectateur sont celui de l'intertexte pour le pôle humain, et celui du «savoir de l'arché» pour le pôle machinique. Cette conception bipolaire conduit parfois, j'en conviens, à certaines simplifications. Par exemple, le «savoir de l'arché» porte aussi sur l'intervention humaine dans la genèse du film, et l'intertexte peut, lui aussi, amener l'audiospectateur à percevoir, par comparaison, un marquage du signifiant filmique. Toutefois, je me contente de fournir ici une vue synthétique des points abordés en faisant figurer des processus dominants.

En ce qui concerne l'image, j'ai également pris en compte une manifestation verbale peu traitée dans ce travail, puisque mon objectif consistait à aborder les phénomènes d'oralité: les mentions écrites. Certaines d'entre elles sont comparables à la voix-over du point de vue de leur organisation discursive et de leur contribution à l'élaboration du discours filmique (c'est pourquoi elles renvoient également à une instance fictive), sous toutefois bénéficier de la personnalisation induite par la vocalité.

On constate qu'il y a deux niveaux de personnification, c'est-à-dire de constitution d'un sujet parlant: le premier se rapporte exclusivement à l'univers diégétique, tandis que le second articule cette humanisation avec l'instauration d'une instance fictive de méga-énonciation. Par exemple, la voix synchrone d'un présentateur scénique intégré dans un film contribue à particulariser cet individu fictionnel, mais n'est pas associée à l'origine du discours filmique dans son ensemble. En fait, le maillon qui permet de passer de la première personnification à la seconde est la voix-over, dont la singularité réside dans la place fictive qui lui est octroyée au sein de la structure de l'énonciation filmique.

Ce modèle énonciatif de la représentation audiovisuelle (qui n'inclut donc pas le bonimenteur, fondamentalement extérieur à celle-ci) est le point d'aboutissement de cette étude dans laquelle j'ai tenté de combiner une approche narratologique avec la prise en compte de facteurs culturels et techniques propres à un spectacle qui conjugue une dimension machinique et des manifestations vocales indéfectiblement associées à l'humain. Ce chapitre VI n'a toutefois été consacré qu'à l'un des pôles du «dispositif cinéma», celui de la représentation. On retrouve néanmoins l'ingérence inévitable des deux autres pôles: le spectateur (qui établit par inférences le degré de marquage énonciatif et de personnalisation des instances) et la machinerie (à laquelle renvoie, dans la représentation même, le marquage matériel de l'énonciation). Comme je n'ai traité du personnage qu'en lien avec la création fictive d'un «méga-énonciateur», j'ai fait somme toute peu de cas des questions de «focalisation» dont traitent la plupart des ouvrages consacrés à l'énonciation filmique. J'ai préféré limiter les effets de celle-ci à un niveau second (celui de la diégèse et de l'identification du spectateur au personnage) afin de ne pas occulter le travail du (et sur le) signifiant filmique. En ce sens, mon approche tente de renouveler la question de l'énonciation au cinéma en mettant

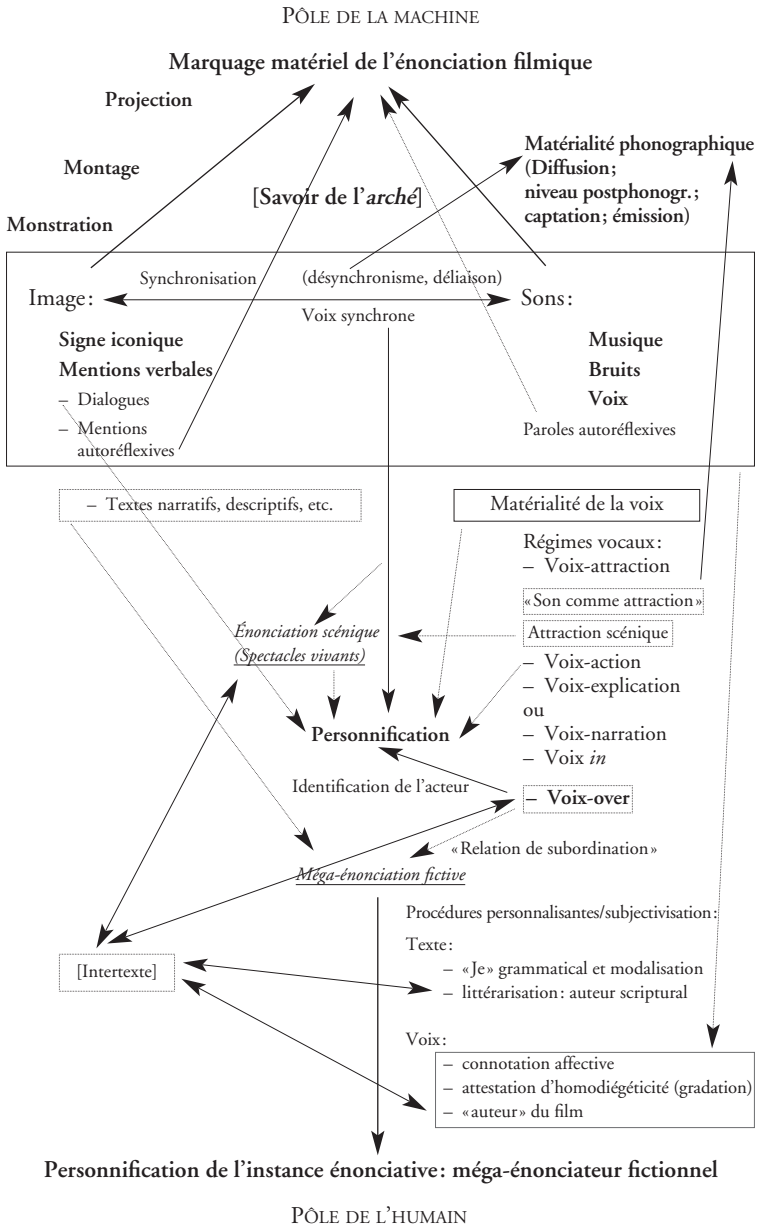
l'accent sur la dimension sonore et, plus généralement, machinique du médium, rejoignant ainsi une proposition émise par André Gaudreault dans un article de 1997 où, après avoir fait état des nombreuses ambiguïtés soulevées par l'application aux textes scripturaux et filmiques de la narratologie (post)genettienne, le narratologue esquissait en ces termes les contours d'une nouvelle voie pour la théorie de l'énonciation (narrative) au cinéma :

« Il faudrait traiter toute la trame narrative englobante qu'est le substrat audiovisuel d'une tout autre manière. Ainsi arriverait-on, peut-être, à imaginer d'autres concepts que ceux de narration et de focalisation pour ce récit qui se fait et se livre par l'intermédiaire de ce dispositif machinique qu'est le cinéma et qui est au fond tout à l'opposé de ce récit verbal, écrit ou non, dont la première évidence est ce caractère humain, plutôt que machinique, qui lui vient de sa nature proprement linguistique et langagière. »<sup>774</sup>

Le modèle d'énonciation que j'ai développé ici tente de considérer le caractère machinique évoqué par Gaudreault dans ce texte (ainsi que, çà et là, dans son ouvrage *Du littéraire au filmique*) et, comme il le propose, de prendre la mesure des implications théoriques liées à la spécificité des dispositifs audiovisuels, « matériellement » fort différents des textes scripturaux. Toutefois, mon approche ne s'est ni cantonnée à l'organisation narrative, ni appuyée sur une opposition entre le « dispositif machinique » et le verbal, puisque ce dernier représente selon moi une composante essentielle de la communication filmique, et cela même à l'époque du cinéma dit « muet ». Au contraire, la question des relations entre paroles et images m'a conduit à intégrer la dimension humaine que la « nature langagière » du verbal – et, j'ajouterais, son actualisation dans l'oralité – confèrent à l'appréhension fictive de l'énonciation filmique activée chez l'audiospectateur. Ainsi retrouvons-nous, conceptualisés dans le cadre d'une théorie « pragmatique » de la voix et de la parole au cinéma, l'interaction du destinataire et des deux composantes inséparables du spectacle cinématographique dont j'ai systématiquement tenté de dialectiser la relation : l'humain et la machine.

774. André Gaudreault, « Sistema del relato de un film con narrador verbal: *Citizen Kane* de Orson Welles », *Archivos de la filmoteca*, N° 27, 1997. Je remercie l'auteur de m'avoir fourni la version originale de ce texte, qui n'a pas été traduit en français jusqu'à présent.

Marquage énonciatif et humanisation de la représentation audiovisuelle







## VII. LA VOIX-OVER DANS HIROSHIMA MON AMOUR

### VII.1 LA VOIX (VOIE?) D'UNE CERTAINE MODERNITE

Alain Resnais est l'un des principaux cinéastes pour lesquels la voix revêt une importance primordiale. En observant sa filmographie, on constate un mouvement qui part du pôle de la voix-narration, omniprésente dans ses courts métrages documentaires ainsi que dans ses deux premiers longs métrages *Hiroshima mon amour* (1959) et *L'Année dernière à Marienbad* (1961) (on peut omettre *Ouvert pour cause d'inventaire*, film 16 mm qui n'a pas connu de sortie publique), pour aboutir à la voix-attraction de ses deux dernières réalisations (*On connaît la chanson*, 1997, et *Pas sur la bouche*, 2003). Dans l'intervalle, il exploite divers usages de la déliaison vocale, par exemple dans *Muriel* (1963), film pour lequel il dit s'être inspiré de la bande dessinée pour créer des effets de chevauchement du texte sur les unités visuelles<sup>775</sup>, ou dans des films «mixtes» en termes de régimes vocaux qui intègrent soit des parties chantées (*La Vie est un roman*, 1983), soit des voix provenant d'un «monde possible» parallèle (l'univers des cartoons dans *I Want to Go Home*, 1989). Précisons que Resnais recourt systématiquement à la postsynchronisation afin de disposer d'une liberté totale en ce qui concerne la détermination des rapports entre la parole et l'image. Dans *Hiroshima mon amour*, film à voix-over dont l'étude me permettra d'appliquer et de discuter certaines propositions avancées au cours du chapitre VI, cette technique s'est imposée comme une nécessité, l'acteur japonais Eiji Okada ne parlant pas du tout français au moment du tournage.

En raison de leur complexité narrative et formelle, les films d'Alain Resnais constituent un objet d'étude privilégié pour les théoriciens, *Muriel* ayant par exemple donné lieu à un ouvrage collectif

775. Cité par François Thomas, *L'Atelier Alain Resnais*, Paris: Flammarion, 1990, p. 241.

résultant d'une analyse d'inspiration structuraliste extrêmement fouillée. Cette recherche collective comporte notamment un chapitre intitulé «Un film sonore, musical, parlant», dans lequel il est question du rôle que joue la parole dans la structure discursive du film<sup>776</sup>. Par ailleurs, Marie-Claire Ropars, qui collabora à cette publication, a développé par la suite une réflexion autour de la notion d'«écriture» qui doit beaucoup à la conjonction originale du texte et de l'image dont font preuve les films de Resnais, et principalement *Hiroshima mon amour*<sup>777</sup>. Les premières réalisations de Resnais se caractérisent effectivement par l'exhibition d'une littérarité qui est le fruit d'une collaboration particulière avec des auteurs dont les textes, écrits spécifiquement pour les films, sont scrupuleusement respectés lorsqu'ils sont proférés par les acteurs. À la sortie d'*Hiroshima mon amour*, Louis Seguin disait de son réalisateur qu'il «a eu le courage d'abolir entièrement les frontières factices qui existaient entre le cinéma et la littérature»<sup>778</sup>. Si l'étiquette de «nouveau cinéma», utilisée en 1966 par Bernard Pingaud à propos du cinéma de Resnais et de Robbe-Grillet<sup>779</sup>, ne correspond aucunement à la constitution d'un mouvement (ou même d'une «famille») comparable au «Nouveau roman» de la littérature – sur le caractère problématique de la constitution de ces deux «mouvements» et sur les liens que l'on peut néanmoins tisser entre certaines pratiques littéraires et filmiques, voir l'ouvrage de Claude Murcia<sup>780</sup>, force est de constater, dans la France du début des années 70 (et même plus tôt pour Agnès Varda, véritable précurseur), l'émergence d'un cinéma qui se distingue foncièrement de la Nouvelle Vague dans ses sources d'inspiration et son style, et présente une parenté avec une certaine modernité littéraire. Resnais a joué en quelque sorte un rôle de passeur entre les deux moyens d'expression en éveillant chez Marguerite Duras et Alain Robbe-Grillet un intérêt pour la réalisation cinématographique, et peut-être en prouvant qu'un film non

776. Michel Marie, «Un film sonore, musical, parlant», in Claude Baiblé, Michel Marie et Marie-Claire Ropars-Wuilleumier (dir.), *Muriel. Histoire d'une recherche*, Paris: Éditions Galilée, 1974, pp. 61-122.

777. Voir Marie-Claire Ropars-Wuilleumier, *Écraniques*, Lille: Presses Universitaires de Lille, 1990, p. 35.

778. «Quoi de neuf? Débat», *Positif*, N° 31, novembre 1959, repris in Stéphane Goudet (dir.), *Positif, revue de cinéma. Alain Resnais*, Paris: Gallimard, 2002, p. 56.

779. Bernard Pingaud, «Nouveau roman et nouveau cinéma», *Cahiers du cinéma*, N° 185, 1966, p. 27.

780. Claude Murcia, *Nouveau roman, nouveau cinéma*, Paris: Nathan, 1998.

dominé par le réalisme de la voix synchrone et le primat de la voix-action était concevable dans l'économie du cinéma de l'époque. Si l'on considère ces passerelles entre cinéma et littérature, *Hiroshima mon amour*, première fiction du cinéaste qui fit beaucoup parler d'elle en raison de la controverse soulevée par son prix au Festival de Cannes, me paraît faire date dans l'«histoire de la voix au cinéma». Un compte-rendu des *Cahiers du Cinéma* se faisait l'écho, à la sortie du film, du trouble provoqué par l'usage non conventionnel qui y est fait de la parole et de la voix :

«Une autre objection que l'on entend fréquemment est que le film est bavard et que le texte de Marguerite Duras est prétentieux et amphigourique. Il n'y a rien à répondre à ceux qui prétendent ne pas aimer ce texte, sinon qu'on l'aime. Mais on peut faire quelques remarques sur l'utilisation du langage et des mots dans *Hiroshima*, parce que c'est, en effet, un des aspects les plus déconcertants du film. Ces remarques valent en même temps pour la façon dont ce texte est dit, qui en est inséparable.»<sup>781</sup>

C'est donc autant la nature du texte que son actualisation vocale proprement dite que d'aucuns ont considérées comme gênantes. Le numéro du mois suivant de la même revue, qui comprend une image du film de Resnais en couverture, propose la retranscription d'une table ronde réalisée autour de ce film par les principaux (futurs) représentants de la Nouvelle Vague. En dépit de leur admiration pour *Hiroshima mon amour*, ceux-ci paraissent quelque peu embarrassés, surtout Éric Rohmer qui affirme à la fois que Resnais «est le premier cinéaste moderne du cinéma parlant» et que son film provoque une certaine gêne, un agacement. Les intervenants font invariablement surgir l'opposition classique *vs.* moderne, tout en témoignant d'une difficulté à classer cette météorite, à l'inscrire dans une filiation<sup>782</sup>.

Le caractère novateur d'*Hiroshima mon amour* tient surtout à la façon dont ce film exploite le potentiel créatif des phénomènes de (co)référence verbo-iconique. Alain Resnais, monteur de formation et principalement préoccupé par des questions de structure, propose une organisation complexe des relations entre les mots et

781. R. G., «Hiroshima mon amour», *Cahiers du Cinéma*, N° 96, juin 1959, p. 10.

782. «Hiroshima, notre amour» (table ronde), *Cahiers du Cinéma*, N° 97, juillet 1959.

les images. Comme on le verra, son film présente trois modalités d'utilisation de la voix-over qui ont des conséquences différentes sur la perception de l'origine fictive de l'énonciation filmique, notamment en ce qui concerne la question de la subjectivisation. L'usage protéiforme de la voix-over proposé par Resnais dans *Hiroshima mon amour* nous permet de discuter les fondements de ce procédé tout en saisissant concrètement quels enjeux il peut soulever dans l'économie textuelle d'un film donné.

## VII.2 TROIS REGIMES OVER: LA CONSTRUCTION PROGRESSIVE D'UN SUJET (PARLANT) FICTIONNEL

Le texte *over* est distribué dans *Hiroshima mon amour* sur trois moments, et ne se manifeste jamais en dehors de ceux-ci. Il s'agit du prologue, du récit en *flash-back* de Nevers et des pensées qui assaillent la Française lorsqu'elle chemine, suivie à distance par son amant, dans les rues d'Hiroshima au cours de la dernière nuit qu'elle passe dans cette ville. Pour chacun de ces segments, une fonction spécifique est impartie à la voix-over: l'instauration de l'univers filmique, l'enchâssement narratif et le monologue intérieur. A priori, ces usages de la voix-over ne paraissent donc pas instituer de rupture avec les pratiques traditionnelles, puisqu'il s'agit là, comme nous l'avons vu, des principales fonctions conférées à ce procédé depuis la fin des années 30. Il faut néanmoins prendre en compte les caractéristiques du texte, des voix et de la visualisation des locuteurs pour saisir les particularités du film de Resnais. La succession de ces trois types de voix-over décrit un mouvement général d'intériorisation sur le personnage de la Française dont l'identité se constitue peu à peu, d'abord à travers la façon dont elle présente le drame d'Hiroshima, ensuite grâce à la médiation de son auditeur diégétique, enfin par introspection.

Les images relatives au bombardement atomique qui composent le prologue d'*Hiroshima mon amour* sur une durée de 13 minutes 15 secondes sont constamment accompagnées d'une voix-over (à l'exclusion de toute autre voix), celle d'Emmanuelle Riva qui interprète la Française ou celle d'Eiji Okada qui lui répond (les personnages n'ont pas de nom, le film se terminant sur l'assignation mutuelle d'un nom de ville qui replonge ceux-ci dans l'anonymat). La voix-over de Riva permet de rassembler dans cette partie inaugurale un matériau visuel hétérogène en lui conférant

une continuité et en précisant sa signification. Ce prologue présente en effet un caractère très fragmentaire, puisqu'il contient 132 des 423 plans du film<sup>783</sup>, alors qu'il s'étend sur moins d'un huitième de sa durée totale. Entremêlant archives, huit plans extraits d'une fiction (*Hiroshima*, Hideao Sekigawa, 1953)<sup>784</sup> et des images tournées par l'équipe de Resnais à Hiroshima, le début de ce film ressemble fort au travail effectué par le cinéaste pour *Nuit et brouillard*. Tout comme ce documentaire sur les camps de la mort nazis, le prologue d'*Hiroshima mon amour* aborde la question du caractère irreprésentable de l'horreur, mais en dépassant cette impossibilité grâce au basculement dans la fiction intimiste de Duras. La facture de cette introduction s'explique par la genèse du film : à la demande du producteur Anatole Dauman d'Argos Film (société qui avait déjà produit *Nuit et brouillard*), désireux de profiter d'une opportunité de cofinancement avec le Japon, Resnais s'est attelé à un projet de documentaire sur Hiroshima. Il y travaille avec Chris Marker, puis décide de renoncer, jugeant qu'il n'est pas possible de se démarquer significativement des films déjà réalisés sur le sujet<sup>785</sup>. L'intertexte composé des autres films du cinéaste traitant des atrocités qui ont bouleversé la conscience de l'homme du XX<sup>e</sup> siècle (surtout *Guernica* et *Nuit et brouillard*) est important pour évaluer comment ces images liminaires étaient reçues à l'époque. Resnais n'ayant réalisé jusque-là que des films non fictionnels de format court ou moyen – la durée de ce prologue correspond approximativement à celle du film *Guernica* – cette entrée dans *Hiroshima mon amour* évoquait nécessairement au spectateur de la fin des années 50 les productions documentaires antérieures du cinéaste. Toutefois, les données paratextuelles et les conditions de projection indiquaient concurremment que le film s'inscrivait dans l'institution du film de fiction. Cette double orientation contractuelle crée une tension qui se reflète également dans l'usage hybride de la voix-over, située à mi-chemin entre le commentaire d'un documentaire et l'expression de la subjectivité d'un personnage de fiction. Pour signifier que la réalité, insaisissable, échappe à toute visualisation, le verbal est d'un grand

783. Pour la référence aux plans, je me baserai sur le découpage proposé par Raymond Ravar, *Tu n'as rien vu à Hiroshima*, Bruxelles : Éditions de l'Institut de Sociologie, 1962, pp. 249-306.

784. D'après Jean-Louis Leutrat, *Hiroshima mon amour*, Paris : Nathan, 1994, p. 23.

785. Laure Adler, *Marguerite Duras*, Paris : Gallimard, 1998, p. 511.

secours: sa nature sémiotique et ses capacités réflexives lui permettent d'exprimer l'« indicible » et de désamorcer la valeur indiciaire de l'image ciné-photographique. À travers le démenti constant du Japonais, ce prologue exploite pleinement le potentiel de la négation linguistique, dont le médium cinéma est privé en raison de l'absence de la « double articulation » dont disposent les langues naturelles.

L'ouverture du film est d'un intérêt tout particulier pour l'examen d'aspects dont il a été question au chapitre précédent. Cette séquence contribue à poser l'univers diégétique et à définir le contrat de lecture, et génère une situation de déliaison en montrant des corps tout en occultant les visages. De plus, la mise en place de la situation dialogique propose un intermédiaire entre l'intériorisation de la parole et la communication diégétique, c'est-à-dire entre un discours « égocentrique » et une manifestation extériorisée. Enfin, l'interpénétration du documentaire et de la fiction permet d'examiner l'amorce d'un phénomène de subjectivisation que le film accentuera progressivement. Étant donné son statut hybride, ce prologue se prête avantagement à l'application de paramètres que nous avons conçus de façon graduelle, tels que la fictionalité, la subjectivité, la déliaison ou le degré de présence d'un certain régime vocal.

Situé entre les minutes 43 et 59 du film (plans 269-323), le deuxième segment ne comporte quant à lui pas continûment une voix-over: sur les 54 plans considérés, seuls 22 en présentent une, alors que sept autres, visualisant le récit enchâssé, en sont dépourvus. Les 25 plans restants sont consacrés à la situation d'énonciation, c'est-à-dire à la discussion entre la Française et le Japonais dans le tea-room, véritable « scène » de la discussion. La place considérable qui est allouée au récit-cadre grâce au principe de l'aller-retour souligne l'importance que prennent les liens entre passé et présent – le second n'étant que la répétition du premier – dès lors que se déclenche chez la Française le processus cathartique de remémoration d'un vécu refoulé. La relation qui s'instaure alors entre la locutrice et son auditeur s'apparente à une cure psychanalytique<sup>786</sup> dans laquelle la verbalisation se fait l'adjuvant d'une guérison, pour autant que le soliloque ne conduise pas au délire (la

786. Voir Bernard Pingaud, « À propos d'*Hiroshima mon amour* », *Positif*, N° 35, 1960, repris in S. Goudet (dir.), *Positif, revue de cinéma. Alain Resnais, op. cit.*, p. 74.

gifle donnée par le Japonais semble empêcher que la narratrice ne rechute dans sa « folie » d'autrefois). Cette comparaison révèle la nature singulière de cet « échange » quasiment dépourvu de réciprocité, et montre combien le monologue désactive la « voix-action » : la parole surgit comme un flux presque ininterrompu adressé à une personne située « hors-champ », à l'instar de l'analyste qui, dans le modèle freudien, est assis en retrait, et donc ne fait pas face au locuteur comme dans une véritable situation de dialogue. Ainsi, quand les voix deviennent *over*, le Japonais disparaît en tant que présence concrète, s'efface au profit de la visualisation (toujours muette) des représentations mentales de la « patiente », dans lesquelles il est parfois intégré sous la forme d'un « tu ».

À mon avis, on peut aussi considérer (cette seconde hypothèse n'exclut pas celle de la « cure ») cette communication verbale majoritairement unidirectionnelle comme une mise en scène de la « maïeutique » resnaisienne qui caractérise le rapport instauré par le cinéaste avec l'écrivain. À l'image du film qui conjugue la parole de Duras et les images de Resnais, l'histoire naît de l'interaction verbale entre un « je » féminin et un « tu » masculin. La thématique du film suggère d'ailleurs ce rapprochement : comme les voix, les couples se lient et se délient. Or la voix du film, c'est avant tout celle de Duras elle-même, puisque le cinéaste lui demandait d'enregistrer sur cassette la lecture de son propre texte, qu'il écoutait ensuite au Japon, ainsi qu'il l'explique rétrospectivement à l'écrivain : « Je n'oublierai pas mes étranges journées passées dans ma chambre en compagnie de votre voix et de deux poupées articulées en bois chargées de remplacer Okada et Riva. »<sup>787</sup> Ainsi les acteurs du film viendront incarner des « machines parlantes » utilisées comme supports d'une voix unique, celle de l'écrivain. Rarement la conjonction de la « voix narrative » (du style propre à un auteur) et de la voix « physique » ne s'est réalisée aussi pleinement. Car si Resnais s'est intéressé au travail de Marguerite Duras, c'est en vertu des qualités rythmiques et acoustiques de son écriture : « Je venais de voir *Le Square* et j'avais été ému par la musicalité de la langue... Et puis j'adorais *Conversation en Sicile* de Vittorini [proche ami de Duras qui n'a pas été sans influence sur son écriture] que je traduisais chaque soir à des copains »<sup>788</sup>. Comme

787. Cité par L. Adler, *op. cit.*, p. 519.

788. *Ibid.*, p. 512.

l'année suivante lorsqu'il travaillera avec Robbe-Grillet, ce sont les traits prosodiques qui touchent Resnais, notamment sous la forme d'une scansion répétitive qui donne au texte l'allure d'une incantation poétique plus que d'un récit. C'est à travers son actualisation dans la matière vocale que la «voix narrative» connaît dans *Hiroshima mon amour* un renouvellement — car le récit est par ailleurs assez rudimentaire, renouvellement qui est aussi celui de l'énonciation filmique (fictive) proprement dite. Resnais parle du ton des répliques du *Square* en termes d'émotion : il me semble que c'est là une dimension qu'il ne faut pas perdre de vue lorsqu'on discute l'impact des voix-over d'*Hiroshima mon amour*, qui agissent à mon sens comme un facteur d'immersion diégétique. L'«artificialité» présumée du texte est en fait reversée sur les attentes nourries envers une «œuvre d'art», et subsumée par des composantes «sensorielles» véhiculées tant par le contenu des paroles que par la matérialité vocale.

En ce sens, en dépit de la volubilité de ses personnages, Alain Resnais, qui souligne fréquemment dans ses entretiens la déception ressentie enfant lors de la généralisation du parlant, renoue inconsciemment avec un idéal «musicaliste» des années 20 qui n'est probablement pas étranger à sa formation de monteur : le rythme géré à travers la synchronisation vocale de mise en chaîne constitue un élément-clé de son esthétique. Nous ne sommes donc pas surpris d'apprendre dans un entretien récent que Resnais tenait à accompagner les projections 16 mm qu'il organisait chez lui en synchronisant manuellement tourne-disques et images<sup>789</sup>. À propos des projections privées de son premier long métrage, *Ouvert pour cause d'inventaire*, il déclare de même : «Il y avait bien trente ou quarante plages de disques différents pendant la projection. Il fallait synchroniser et je sortais de là épuisé.»<sup>790</sup> On retrouve chez Resnais ce goût de la synchronisation *live* telle que la pratiquèrent tant de pionniers de l'époque «muette». Le rythme insufflé par la musique constitue une composante essentielle de ses films. Fréquemment, il compare son travail à celui d'un compositeur<sup>791</sup>. Il y a bien chez lui une compréhension «métaphorique» de la musique qui rejoint le cadre théorique de cinéastes des années 20 comme Gance ou Dulac. C'est pourquoi la musique de

789. S. Liandrat-Guigues et J.-L. Leutrat, *op. cit.*, p. 177.

790. *Ibid.*, p. 186.

791. Voir par exemple B. Pingaud, «Alain Resnais», *Premier Plan*, N° 18, 1961, pp. 55-56.



Fusco est si présente dans *Hiroshima mon amour* (et même ininterrompue sur l'ensemble du prologue) où elle contribue de manière décisive, par le jeu des thèmes qu'elle convoque, à l'organisation narrative (voir à ce propos l'analyse d'Henri Colpi)<sup>792</sup>. Il est intéressant de constater que, pour le cinéaste, la « musicalité » du film se situe en premier lieu au niveau de la voix des acteurs et dans le ton du texte. En cela, il annonce la modernité cinématographique incarnée par Jean-Marie Straub et Danièle Huillet (dont le premier film date de 1962) qui, justement, ont récemment recouru à la prose vittorinienne (leur film *Sicilia!* de 1998 est réalisé d'après la *Conversazione in Sicilia*) à laquelle Resnais fait référence dans l'entretien cité ci-dessus.

Le deuxième segment en voix-over, d'une durée à peu près équivalente à celle du prologue, repose sur le principe traditionnel de l'alternance d'une situation-cadre d'énonciation et de la visualisation d'un enchâssement. Pour trois raisons, ce *flash-back* se distingue du modèle plus classique que représente par exemple *Le Jour se lève* (Marcel Carné, 1939) : à un niveau macrostructurel, son extension considérable implique une disjonction par rapport au premier « *flash-back* » exposé dans le prologue, où la grande histoire précède l'« histoire de quatre sous » ; ensuite, cet enchâssement est faiblement motivé sur le plan narratif dans le récit-cadre, si ce n'est par le fait que le Japonais veut obtenir de la Française une sorte de « don virginal » dans la confession ; enfin, cette histoire est présentée de façon morcelée et achronologique<sup>793</sup>. L'aspect que Yannick Mouren considère comme véritablement novateur dans le traitement des *flash-backs* d'*Hiroshima mon amour* et qui a fait moult émules – il est totalement banalisé aujourd'hui – consiste en l'irruption inattendue d'une association mémorielle, comme celle qui a lieu lorsque la Française, contemplant son amant étendu sur le lit, est assaillie par un flash montrant le soldat allemand couché sur le sol dans une position similaire<sup>794</sup>. Le principe d'incertitude qui sous-tend cet insert isolé, appel du passé et source d'une énigme pour le spectateur puisque le plan est trop bref pour être décrypté, est également à l'œuvre par la suite dans le récit *over*, certaines images ne trouvant aucune explication

792. Henri Colpi, « Musique d'Hiroshima », *Cahiers du Cinéma*, N° 103, 1960, pp. 1-14.

793. Sur cette question de l'« ordre » dans le récit du film, voir René Prédal, *Études cinématographiques*, N° 64-68 (« Alain Resnais »), 1968, p. 138.

794. Yannick Mouren, *Le Flash-back*, Paris : Armand Colin, 2005, p. 138.

immédiate dans le verbal, ou n'étant accompagnées d'aucune parole. Ces lacunes provisoires exhibent l'un des principes-clés de l'organisation signifiante du film : le retardement de la coréférence verbo-iconique. Resnais joue fréquemment sur le « suspense structurel » pour garder l'audiospectateur en éveil ; on a vu que l'usage de ce procédé dans *Hiroshima mon amour* avait frappé Iouri Lotman (voir chapitre IV. Intégration musicale de la parole).

D'ailleurs, la voix-over n'apparaît pas tout de suite pour commenter les images de Nevers. Dans une première série de plans (223-255) consacrés au passé de la Française, les segments de *flash-backs* sont préalablement introduits verbalement en voix *in*, puis uniquement accompagnés de musique. La manifestation *over* me semble indiquer que la « patiente » est prête à se dévoiler à travers son récit, à faire sortir ce qu'elle renferme dans ses « profondeurs », dont l'obscur souterrain de Nevers se fait l'expression métaphorique. C'est significativement un plan de la cave qui ouvre cette partie et permet le basculement dans l'enchâssement : « La cave est petite... » (voix *in*), « ...très petite » (voix-over), dit la Française, comme si les dimensions restreintes de cet endroit étaient à l'image du retranchement dans l'intimité d'un monde intérieur. Notons que cette narration *over* est aménagée progressivement : dans le premier plan parlant de la scène située dans le tea-room du bord du fleuve, alors que la Française débute un récit dont elle tente d'abord de nier le caractère personnalisé (« Nevers : 40 000 habitants ; bâtie comme une capitale, ... »), le visage de Riva est caché au niveau de la bouche par celui du Japonais filmé de profil, situation *off* qui fait office d'amorce pour le texte *over*. Ensuite, nous glissons dans l'histoire de l'amour interdit de Nevers au moment de l'énonciation du mot « douleur », immédiatement suivi d'un bref plan montrant le soldat allemand agonisant sur la berge (un contenu dont le spectateur ne peut avoir qu'une compréhension rétrospective), réminiscence fugace qui n'est cependant accompagnée d'aucune voix. Ce n'est qu'à partir de la description *over* de la cave que la Française semble véritablement prendre en charge son récit, instaurant passagèrement une relation de subordination de l'image au texte que les percées de flashes énigmatiques viennent ébranler. Lorsque la Française affirme que « Les mains deviennent inutiles dans les caves », le référent verbal se rapporte autant à la situation présente – elle serre les mains de son interlocuteur – qu'à l'enchâssement, il est le lieu de la bifurcation du plan de référence.

Durant le récit de la période de folie, les indications temporelles du verbal sont essentielles pour dilater le temps de l'histoire, car les plans enchâssés sont généralement assez brefs. Le texte peut présenter un circonstant temporel précis («à six heures du soir», «l'hiver est terminé»), mais le plus souvent il confère une valeur itérative («Je la vois marcher cette société. Rapidement pendant la semaine. Le dimanche, lentement», «Quelquefois un chat rentre et regarde», «Des gens passent sur le pont, plus ou moins nombreux suivant les heures»). Cette plongée dans l'abîme d'un souvenir traumatique provoque nécessairement une forte subjectivisation du récit de la narratrice.

Tout le film est centré sur la constitution de ce personnage dont l'intériorité nous est livrée sans aucune médiation dans le dernier des trois segments *over*, qui débute au plan 351 (à 1:07:54) lorsque la Française, qui a regagné sa chambre d'hôtel, se rafraîchit le visage avec l'eau du lavabo. Le monologue intérieur surgit juste après l'image d'un miroir dans lequel la protagoniste se regarde: comme dans la première séquence de *Cléo de 5 à 7* (Agnès Varda, 1961), c'est une «réflexion» qui introduit le soliloque. Le spectateur est ainsi mis dans les mêmes conditions de perception visuelle que le personnage, puisque la Française, elle aussi, se voit «de l'extérieur». Cette analogie de positionnement, qui participe de la construction de soi (une sorte de «stade du miroir»), me paraît être l'amorce d'un basculement dans l'intériorité. Il s'agit, selon notre typologie des rapports entre voix-over et image, d'un cas de «dissociation»: l'actrice ne remue pas les lèvres, mais on peut entendre ce que pense son personnage. Un désancrage plus important s'opère par la suite, après le premier des deux dialogues *in* qui ponctuent le flux *over* de ce segment. En effet, des plans de Nevers et d'Hiroshima alternent dans un grand maelström où les espaces se confondent. Le temps verbal du présent, systématiquement utilisé par la narratrice pour évoquer les événements de Nevers, trouve dans ce mélange d'images un équivalent visuel de la présentification du passé. On renoue en quelque sorte avec la situation *over* du début où la locutrice semblait se projeter dans les espaces visualisés, mais cette fois en intégrant Nevers dans Hiroshima, comme si la distinction entre récits enchâssant et enchâssé avait été abolie. L'écho instauré avec le prologue, tant au niveau du texte («Tu me tues...») que de la musique – Henri Colpi mentionne à propos de ce passage l'«intervention de l'orchestre qui rappelle les

principaux thèmes du Prologue»<sup>795</sup> – crée un effet de clôture. La situation est néanmoins différente du début du film, puisque désormais le spectateur est familier de la psychologie du personnage. Il me paraît donc important de revenir aux premières images du film, où l'origine du discours n'avait pas encore ce caractère d'évidence.

### VII.3 LE PROLOGUE: FUSION DES CORPS, FISSION AUDIOVISUELLE

Nous avons constaté à plusieurs reprises combien la voix-over est prédisposée à assurer l'entrée du spectateur dans le monde du film. C'est pourquoi j'aimerais analyser son rôle dans le prologue d'*Hiroshima mon amour*, qui est à cet égard passablement singulier, et même totalement fascinant. Il s'ouvre sur une série d'images ne montrant aucun visage, mais seulement des corps enlacés, filmés en plan rapproché et fragmentés par le cadre. La parole n'a donc pas de lieu d'émission dans l'image, mais pourrait en avoir un: il y a indétermination de l'ancrage diégétique. Le fait que l'on commence par montrer les corps semble suggérer l'ancrage (voix *off*); par contre, la qualité de l'enregistrement, dénué de toutes traces de bruits ambiants, ainsi que la matérialité vocale, qui paraît ne renvoyer en rien au contexte physique dont elle est censée provenir – les acteurs s'adonnent ostensiblement à un acte de lecture, ils ne miment pas une discussion qui se déroulerait lors d'une étreinte – donnent l'impression que ces voix se détachent du vide, émanent d'un «hors-film». «Le personnage dans le cinéma de Resnais [...] est passé par la mort et il naît de la mort», affirmait Deleuze<sup>796</sup>. Dans *Hiroshima mon amour*, ces êtres d'outre-tombe – caractère quasi surnaturel souvent représenté au cinéma par des voix déliées – nous apparaissent sous une forme désincarnée. Néanmoins, dans leur matérialité même, les voix donnent déjà quelque chose à connaître de la diégèse qui laisse présager l'homodiégéticité ou l'ancrage: l'opposition masculin *vs.* féminin et francophone *vs.* asiatique se perçoit immédiatement, et peut correspondre aux deux corps visualisés. Une autre dichotomie permet d'esquisser un premier «portrait» des personnages: à la volubilité et à l'autojustification de la femme répondent le laconisme et l'assurance de

795. H. Colpi, *art. cit.*, p. 10.

796. Gilles Deleuze, *Cinéma II. L'Image-temps*, Paris: Minuit, 1985, p. 270.

l'homme, que l'accent lie au passé d'Hiroshima dont il est question. Alors qu'un monologue accrédi-terait le statut *over* de la voix, le dialogue tend à ancrer la parole dans les deux corps. Or le spectateur est livré à une indétermination qui renforce l'impression de «flottement» propre à cette partie inaugurale: un commentateur émet même l'hypothèse selon laquelle la voix masculine ne serait rien d'autre que l'expression verbalisée des scrupules qui tourmentent la Française, dont nous partagerions dès lors exclusivement les pensées<sup>797</sup>.

Toutefois, au niveau du «synchronisme», l'impression d'une déliaison subsiste qui s'inscrit dans une dialectique de la fusion (les objets d'Hiroshima fondus par la chaleur extrême de la bombe: «dix mille degrés sur la place de la Paix») et de la fission (celle de l'atome, à laquelle correspond la fragmentation des corps par le cadrage). À l'union physique consommée dans l'amour ou dans la mort (les corps nus se couvrent successivement de cendres et de sueur) s'opposent la disjonction des contenus verbaux (il nie ce qu'elle affirme) et le maintien d'une séparation entre les voix et l'image. Véronique Campan, décrivant le procédé du doublage, suggère que cette technique engage une fusion: «Une voix, deux fois expropriée, est imposée de l'extérieur à l'image d'un corps dans lequel elle doit se fondre.»<sup>798</sup> En exhibant ce que nous avons appelé la «déliaison fondamentale», le prologue d'*Hiroshima mon amour* enrayer cette fusion, exploite le potentiel dissociatif de la voix postsynchronisée, et donc sa faculté, prônée par Campan, de «dispenser et d'affoler la représentation filmique»<sup>799</sup>. Or il y a dans l'affolement provoqué par ce prologue quelque chose du «Big-bang», d'un chaos originel qui superpose la destruction massive (réfèrent du texte) à la genèse même du film. François Thomas remarquait à propos de l'œuvre de Resnais (à la fin des années 80, car sa remarque ne sera plus valable par la suite): «Tous les films de Resnais naissent du chaos, instaurant dans leurs premières minutes une esthétique du bégaiement [...] avant de s'organiser progressivement [...].»<sup>800</sup> L'étrangeté provoquée par la déliaison vocale, qui

797. Haim Callev, *The Stream of Consciousness in the Films of Alain Resnais*, New York: McGruer, 1997, p. 111, cité in Emma Wilson, *Alain Resnais*, Manchester; New York: Manchester University Press, 2006, p. 50.

798. Véronique Campan, *L'Écoute filmique. Écho du son en image*, Saint-Denis: PUV, 1999, p. 58.

799. *Ibid.*, p. 59.

800. F. Thomas, *L'Atelier...*, *op. cit.*, p. 16.

ne se voit résolue qu'après un quart d'heure de film, participe de cette délitescence liminaire sur laquelle s'édifie graduellement l'identité du personnage de la Française.

Le prologue d'*Hiroshima mon amour* accumule les tensions: entre les corps et les voix puis, lorsque défilent les images d'archives, du musée, etc., entre l'expression vocale d'une subjectivité et la nature objective des documents montrés, ou plus précisément entre l'individu et le collectif. Le sujet parlant ne peut s'approprier une expérience qu'il n'a pas vécue: la matière visuelle le dépasse, et le référent visé est lui-même encore bien au-delà. Le discours assertif est frappé d'inanité dès l'ouverture du film à travers la forme dialogique, puisque la parole du Japonais est première: le déni «Tu n'as rien vu à Hiroshima» précède toute affirmation. En outre, le pronom inaugural «tu» s'adresse aussi au spectateur, l'interlocuteur n'étant pas encore construit sur le plan diégétique. Comme dans le récit ultérieur de Nevers, c'est la voix de l'homme qui se fait d'abord entendre *over*, comme si elle constituait la condition d'émergence de la voix-over féminine. Dans un dialogue réaliste, «incarné», l'audiospectateur interpréterait la phrase «Tu n'as rien vu...» comme un indice de la représentation *in medias res* d'une conversation; ici, elle fait office d'un refus de principe, d'une impossibilité fondamentale qui est liée, en ce qui concerne Resnais, à la réalisation d'un documentaire sur Hiroshima. Le film fictionnel, et donc la voix du personnage de fiction, naissent de cette impossibilité première. L'alternance des voix permet par contre au «je» de la Française de s'affirmer. Je propose de partir du premier énoncé de la locutrice pour tirer, en fonction du membre de phrase sur lequel on fait porter l'accent, une série de conclusions valables pour l'ensemble du prologue.

### **J'AI TOUT VU. TOUT.**

La brièveté de la formule et la répétition du pronom constituent des traits caractéristiques du style durassien, que l'on reconnaît notamment à la simplicité du vocabulaire utilisé et au primat accordé à la dimension prosodique. Cette concision donne à ce premier énoncé une valeur péremptoire soulignée par le ton catégorique d'Emmanuelle Riva, ainsi que par la structure phonétique: d'une part, le mot «tout» est frappé à deux reprises de l'accent dit «tonique» de la phrase française, ce qui est d'autant plus perceptible qu'il est monosyllabique; d'autre part, l'allitération du «t»,

que les phonéticiens définissent dans le système français des consonnes comme sourd et apical, confère une certaine brutalité à l'énoncé, à l'instar de la saccade introduite par la césure du point. Le «t» est en effet classé parmi les sons «durs» qui, selon les résultats des recherches en psycho-phonétique menées par Ivan Fónagy, seraient socialement liés à des «énergies pulsionnelles agressives»<sup>801</sup>. La locutrice est ici sur la défensive, elle évoque ce qu'elle a vécu comme si, mue par un sentiment de culpabilité, elle se sentait contrainte à alléguer des arguments. Dans la suite du prologue, dont le ton assertif contraste fortement avec le récit ultérieur de Nevers, Emmanuelle Riva fait néanmoins sentir qu'un doute s'insinue sous son assurance apparente. Une trop grande précipitation, certains tremblements dans la voix ou une sorte de «flottement» suggèrent des fissures qui sont autant d'échappées vers la psychologisation d'un être de fiction.

Le pronom «tout» condense les référents du discours verbal, instaurant une relation de coréférence avec cet ensemble d'images de natures diverses qui alternent avec les plans sur les corps, dans un mouvement qui actualise l'oxymoron du titre. La promesse d'une complétude de l'univers décrit s'oppose à l'incomplétude de la diégèse visualisée. Ce pronom est également la matrice de la structure énumérative qui, par la suite, permet de mentionner successivement les diverses parties du «tout», et connote la visée globalisante du discours de ce prologue. Dans ses documentaires, et surtout dans *Nuit et brouillard*, Resnais avait déjà usé d'une forme énumérative soutenue par un texte *over* dont la dimension paratactique renvoyait, par sa froideur, à l'inhumanité de la situation représentée (en ce qui concerne par exemple les «styles» architecturaux des camps de la mort ou les préceptes écrits par les nazis sur des pancartes). Or, ici comme là, l'énumération et la catégorisation des faits apparaissent comme le résultat d'une emprise fantasmée sur le réel (et sur le passé) qui, dans le film, se révèle totalement vaine. Par contre, la négation de la représentation est absolue dans le cas d'*Hiroshima mon amour*: alors que, dans *Nuit et brouillard*, Resnais pouvait compter sur le témoignage de Jean Cayrol, ici tout est obtenu de seconde main, et ne résulte donc pas du vécu. Cette différence explique que le texte d'*Hiroshima mon amour*, contrairement à celui du court métrage sur l'univers

801. I. Fónagy, *op. cit.*, pp. 88-95.

concentrationnaire, n'est que très peu explicatif dans sa partie consacrée au bombardement d'Hiroshima.

Alors que *Nuit et brouillard* tentait d'exposer les rouages du fonctionnement abject de la machinerie nazie présentée comme une conséquence de la logique industrielle, *Hiroshima mon amour* expose en effet peu, si ce n'est dans quelques slogans politiques, les relations de cause à effet qui régissent cet événement historique. Dans ce prologue, «tout» y est, mais déversé en vrac dans une logorrhée haletante, annonciatrice du «flux de conscience» qui se manifesterait *over* lorsque le film aura bifurqué de la catastrophe collective aux blessures secrètes de la jeune femme. C'est pourquoi entre le «tout» de la Française et le «rien» du Japonais, il n'y a qu'un tour de parole.

#### **J'AI TOUT VU. TOUT.**

Il me semble que l'acte de vision est thématé dans *Hiroshima mon amour* à partir de l'idée que l'Occident a, littéralement, «fermé les yeux» sur le drame. «Comment aurais-je pu éviter de le voir?», dit la Française à propos de l'hôpital. Sur place, les traces s'imposent au regard et, comme le dit l'héroïne dans la séquence qui suit le prologue lorsqu'elle explique au Japonais pourquoi elle s'est intéressée au passé de la ville bombardée: «De bien regarder, je crois que ça s'apprend.» À cette phrase succède d'ailleurs un fondu au noir, expression de la non-vision qui renvoie le spectateur à sa propre activité perceptive.

Dès sa première manifestation verbale, le personnage de la Française se présente comme un sujet percevant. Ce que le spectateur voit est présenté comme l'équivalent de ce qu'elle prétend avoir vu. Ainsi la dénégation du Japonais porte-t-elle également sur l'acte même qui définit l'activité du spectateur dans la salle. Le désaveu de la valeur d'attestation de la vision s'affiche simultanément comme l'affirmation de la force du verbe, qui, plus tard dans le film, ira jusqu'à s'insinuer sous forme écrite dans l'image (les slogans des pancartes brandies par les habitants qui défilent sont rédigés dans le style de Duras). La relation de subordination semble donc fonctionner en faveur de la parole, reléguant l'image à un rôle d'illustration. Toutefois, le montage, le cadrage et les mouvements d'appareil complexifient cette relation. Au niveau du sens, l'autonomie de l'image est manifeste dans certaines utilisations rhétoriques du contraste sémantique, par exemple lorsque la locutrice évoque des



fleurs alors qu'on nous montre l'atrocité des mutilations. Par contre, l'image se joint à la dénégation verbale dans le plan du panoramique sur le paysage dévasté d'Hiroshima, au moment où la voix masculine demande: «Sur quoi aurais-tu pleuré?» L'image, elle aussi, affirme qu'il n'y a «rien», opposant sa valeur documentaire à la séquence reconstituée du film de fiction qui précède. Dans le cas des images tournées à Hiroshima par Resnais se pose également la question du «qui voit?», autrement dit de ce que François Jost appelle l'«ocularisation». Le prologue écarte certes la structure classique du champ/contrechamp, puisque l'actrice n'est pas visible, mais un regard se dessine en creux dans une série de plans filmés en travelling, tant à l'hôpital que dans le musée. Ce mouvement, qu'il paraît vraisemblable de rapporter au déplacement d'un être humain (la caméra n'est pas fixée au plafond sur un rail, mais longe les couloirs et les objets exposés), suggère une présence. Nous ne voyons par contre jamais la Française à l'image, ce qui interdit de prêter un visage à la locutrice par attestation d'homodiegéticité. Nous sommes dans le cas que Jost qualifie d'«ocularisation interne primaire», où «se marque dans le signifiant la matérialité d'un corps ou la présence d'un œil»<sup>802</sup>. Néanmoins, cette construction est particulière dans ce prologue puisque aucun personnage ne s'est constitué jusque-là, les corps appartenant à un autre espace-temps. De plus, ces travellings s'effectuent dans une régularité qui laisse transparaître leur origine mécanique, court-circuitant ainsi l'attribution fictive d'une origine humaine (fictionnelle) au regard. Christophe Carlier souligne que la vitesse constante du travelling «dit assez qu'il repose sur une supercherie technique»<sup>803</sup>. Le statut ambigu de ce prologue, qui a un pied dans la production documentaire antérieure de Resnais et l'autre dans la fiction où se construit un personnage, influence la façon dont le spectateur comprend le verbe «voir», notamment dans ses rapports au «je». Car, même si la première personne signale l'émergence d'une individualité, l'«individu» demande à se constituer en unité: le processus de la quête identitaire (qui est aussi un «travail de deuil») se déroule en fait sur toute la durée du film.

802. F. Jost, *L'Œil-caméra*, op. cit., pp. 27-28.

803. Christophe Carlier, *Marguerite Duras. Alain Resnais. Hiroshima mon amour*, Paris: PUF, 1994, p. 65.

**J'AI TOUT VU. TOUT.**

Le texte désigne un « je » dont il fait le sujet d'un acte de vision. Ce pronom personnel induit l'homodiégéticité de la narratrice, mais rien ne fait office d'attestation dans l'image : aucun corps de femme n'est individualisé, et a fortiori aucun personnage parlant dont la voix pourrait être reconnue comme étant celle que l'on entend *over*. Sur ce point, il est intéressant de considérer les remarques dont le cinéaste fait part à Duras dans le « carnet de bord » qu'il lui envoyait régulièrement du Japon. Dans une lettre du 3 août 1958, reproduite dans le livret de l'édition DVD Arte du film, Resnais déclare :

« Je me demande encore pour l'instant si on doit montrer le visage de Riva dans le Musée et l'Hôpital, alors qu'on ne l'a pas encore vue dans le lit. La montrer de dos serait encore plus ridicule (on ne s'imagine jamais de dos quand on raconte) et ferait bêtement film policier. Sans doute faudrait-il traiter cette première séquence comme un documentaire. »

La précision donnée entre parenthèses est révélatrice d'une volonté de définir l'enchâssement en fonction d'une attitude propositionnelle de l'énonciatrice verbale qui, néanmoins, n'apparaît pas distinctement sous cette forme dans le film. De toute évidence, Resnais a opté pour le « traitement documentaire », expression que l'on peut comprendre, si l'on considère son travail précédent, au sens de l'évacuation de tout personnage. Le 6 août, jour de commémoration du bombardement, Resnais allègue un facteur plus pragmatique qui témoigne de son attachement à l'ouverture de type « documentaire » :

« Montrer Riva dans l'hôpital et ailleurs peut nous coûter dans les 800 000 yens. Je crois avantageux de transporter cet argent dans la scène du café du fleuve, par exemple, ou dans le défilé. »

Peut-être le prologue d'*Hiroshima mon amour* aurait-il constitué un enchâssement plus traditionnel si le budget de cette coproduction n'avait pas été aussi serré. Il arrive souvent que des contraintes pratiques influencent de façon décisive la facture d'un film. Or la voix-over, procédé peu coûteux, intervient fréquemment dans ces modifications tardives. Si nous avons vu Emmanuelle Riva dès les

premières images d'Hiroshima, la voix n'aurait pas été, pendant près d'un quart d'heure, la seule composante à nous fournir des informations sur le personnage. Dans le film tel qu'il a été réalisé, le grain de la voix de la Française, son timbre, ses intonations prennent une importance considérable. L'ambiguïté de l'ancrage du personnage dans le monde du film renvoie à sa fragilité psychologique. Ce n'est qu'avec la visualisation de l'espace de la chambre d'hôtel et avec la voix synchrone que le monde environnant la locutrice prend véritablement consistance, notamment avec les bruits de rue qui, brusquement, apparaissent sur la piste-son (un surgissement dont l'artificialité dénote le changement de régime discursif). Dans le prologue, le personnage se résume à une double fonction de sujet: *sujet grammatical* de l'énonciation verbale et *sujet d'un regard*.

### **J'AI TOUT VU. TOUT.**

La plupart des phrases rédigées par Duras sont conjuguées au présent ou au passé composé, deux temps qui relèvent, selon Benveniste, de l'énonciation discursive. Ce que le linguiste dit de ce temps verbal me paraît particulièrement intéressant pour analyser le prologue d'*Hiroshima mon amour*, d'autant plus qu'il fait allusion à un pan de la littérature qui est contemporain du film (son article paraît précisément en 1959) et auquel on peut sans nul doute rattacher le travail de Marguerite Duras:

«Inversement, la statistique ferait ressortir la rareté des récits historiques rédigés entièrement au parfait, et montrerait combien le parfait est peu apte à convoquer la relation objective aux événements. Chacun peut le vérifier dans telle œuvre contemporaine où la narration, de parti pris, est entièrement au parfait; il serait intéressant d'analyser les effets de style qui naissent de ce contraste entre le ton du récit, qui se veut objectif, et l'expression employée, le parfait à la 1<sup>re</sup> personne, forme autobiographique par excellence.»<sup>804</sup>

Le passé composé entretient en effet avec le présent de l'énonciation une relation forte dont résulte sa valeur prétéritive. Or on constate que Benveniste discute ce point en se basant sur la

804. Émile Benveniste, «Les relations de temps dans le verbe français» [1959], *art. cit.*, p. 244.

dichotomie objectivité/subjectivité qui joue un rôle essentiel dans *Hiroshima mon amour*: alors que la Française croit dresser un bilan objectif du drame à partir de ce qu'elle a appris, elle parle avant tout d'elle-même, usant de la « forme autobiographique par excellence ». En cela, le prologue annonce l'introspection qui donnera lieu à l'évocation du passé de Nevers.

Notons que le passé composé, s'il est systématiquement lié aux actions des personnages, fait place au couple passé simple/imparfait dans l'une des phrases *over* de ce prologue: « Hiroshima se recouvrit de fleurs. Ce n'étaient partout que bleuets [...] qui renaissaient des cendres avec une extraordinaire vigueur. » La parole est toutefois désignée ici comme une dérive, puisque l'image montre en contrepoint les blessures des irradiés. Le passé simple est utilisé pour sa connotation « littérisante » (déjà ressentie à l'époque, ainsi que l'a discuté Roland Barthes lorsqu'il qualifiait dans les années 50 ce temps verbal de « rituel des Belles-Lettres »)<sup>805</sup>: il renvoie à un type de discours poétique (ou à un langage « fleuri », ainsi que le suggèrent les bleuets) qui vise à magnifier la réalité par la fabrication d'un mythe (celui du Phénix qui renaît de ses cendres), alors que l'écriture de Duras emprunte beaucoup à l'oralité et au langage quotidien. Le passé simple fonctionne donc dans une relation antinomique avec l'énonciation discursive. Le prologue comprend en outre une section où le texte est conjugué au futur, un autre temps qui, pour Benveniste, est propre au discours<sup>806</sup>. Une hypothétique répétition du désastre y est envisagée, alors que l'image montre avec une certaine étrangeté des espaces urbains déserts qui évoquent les films de science-fiction post-apocalyptiques. Après le passé et le futur, l'usage du présent réapparaît à la fin du prologue. Or le référent de ces verbes (« Tu me tues. Tu me fais du bien », « Tu me plais ») a désormais simultanément trait à la situation d'énonciation verbale (le « tu » étant un déictique) et au contexte érotique suggéré par les gros plans sur les corps. Le référent verbal rejoint progressivement les images rapportées aux locuteurs: le « potentiel référentiel » se réduit pour se plier bientôt aux normes de la voix synchrone. Lorsque enfin l'audiospectateur peut mettre un visage sur ces voix, il n'est plus question que des corps: « C'est fou ce que tu as une belle peau. »

805. Roland Barthes, *Le Degré zéro de l'écriture*, Paris: Seuil, 1972 [1953], p. 25.

806. E. Benveniste, « Les relations de temps... », *art. cit.*, p. 243.

Après un détour par l'évocation du désastre atomique, on retrouve le référent de l'«épiderme» – une façon de signifier qu'on retourne à la *surface* des choses – sur l'image de laquelle s'ouvrira le film. Aucune continuité ne s'effectue dans la discussion du couple avec le sujet du «dialogue» *over* qui précède. Cet échange inaugural semble d'abord appartenir à un autre monde, presque à un autre film, comme si un documentaire avait été projeté avant le long métrage, ainsi que la loi l'exigeait en France à partir de 1940<sup>807</sup>. Néanmoins, de nombreux motifs visuels et verbaux du prologue traversent l'ensemble du film et, après quelque temps, le Japonais demande à son amante: «Et pourquoi voulais-tu voir *tout* à Hiroshima?» En dépit de l'absence de continuité verbale, l'affirmation première de la Française a donc bien eu lieu à l'intérieur de l'univers diégétique.

#### VII.4 LES POSSIBILITES DU DIALOGUE

La question de la première personne et du présent m'a conduit à des considérations portant sur l'énonciation verbale; on peut se demander quelle est l'incidence de la modalité du «discours» sur l'énonciation filmique. En effet, dans *Hiroshima mon amour*, la constitution d'un méga-énonciateur fictionnel est retardée. Toutefois, la présence d'une voix-over qui, tout au long du prologue, nous guide en introduisant successivement les différents sujets montrés à l'écran instaure un niveau de méga-énonciation fictive, que le pronom à la première personne et certaines caractéristiques vocales tendent à personnaliser. Cette position de supériorité est parfois contredite provisoirement, mais, dans l'ensemble, le surplomb supposé de l'énonciatrice verbale sur l'image donne l'impression qu'il existe un «foyer» unitaire. En dépit de l'affectation distante que Riva insuffle au film par le truchement de sa diction lyrique (rappelons que cette actrice de théâtre n'avait pas d'expérience de cinéma), l'inscription de l'individualité dans la matérialité vocale s'effectue beaucoup plus fortement dans son cas que dans celui d'Okada. Non seulement la voix de l'actrice, constamment présente, paraît générer l'image, mais Riva est francophone, contrairement à son interlocuteur qui ne s'approprie que

807. Voir Roger Odin (dir.), *L'Âge d'or du documentaire. Europe: Années cinquante*, tome 1, Paris: L'Harmattan, 1998, pp. 19-20.

superficiellement la langue, renforçant l'impression d'acousmatisation. Roland Barthes disait en effet du « grain de la voix » : « [...] cette voix charrie *directement* le symbolique, par-dessus l'intelligible, l'expressif. [...] Le « grain », ce serait cela : la matérialité du corps parlant sa langue maternelle : peut-être la lettre ; presque sûrement la signifiante. »<sup>808</sup> En « parlant sa langue maternelle », Riva charge sa voix d'un au-delà du sens qui participe de la concentration du film sur sa subjectivité ; si l'accent japonais personnalise son interlocuteur, ce n'est que sous la forme générique « un Japonais », non dans l'investissement spécifique d'un « corps » qui travaille mélodiquement la langue en l'imprégnant d'un vécu d'acteur. Par conséquent, le Japonais n'est que très faiblement incarné, demeure quasi abstrait ; on l'assimile plus à une « place » dans la communication qu'à un individu. D'ailleurs, cet aspect touche les deux interlocuteurs, réduits à un *pronom* (« je » / « tu ») qui, littéralement, apparaît à la place du nom : jamais les personnages du film ne sont nommés, si ce n'est dans l'ultime plan où ils endossent le nom de la ville qu'ils « représentent ». Quant à l'autre « tu », l'Allemand interprété par Bernard Fresson, il ne prononce pas une seule parole de tout le film.

Le prologue engage déjà, à mon sens, cette réduction de l'univers diégétique aux deux pôles de la communication à travers l'expression réitérée du plaisir physique : le chiasme « TU ME TU(es) » résume les êtres à leur position abstraite de sujets parlants désincarnés. Cependant, la situation dialogique qui fonde l'organisation du film n'entraîne pas seulement, dans le prologue, une négation du contenu proféré par la voix féminine, mais tend également à happer la voix-over dans une situation de conversation diégétisée. Ainsi la présence même de la voix-over du Japonais met en péril la subordination de l'image aux mots : du monologue au dialogue se joue un possible renversement de la méga-énonciation filmique (fictive) dans une situation plus courante, l'énonciation verbale diégétique (fictionnelle). Le laconisme de l'interlocuteur ainsi que le ton du récitatif situent ce prologue sur le fil du rasoir entre l'ancrage et la déliaison, position hybride qui crée véritablement, comme je l'ai montré à différents niveaux, l'esthétique de cette partie du film. Ce prélude « documentaire » en voix-over ébranle la constitution d'une instance faîtière de l'énonciation filmique pour

808. R. Barthes, « Le grain de la voix », *art. cit.*, p. 238.

mieux mettre en cause la représentation elle-même, jugée impossible. Ce prologue ne joue pas le rôle traditionnel d'*orientation de la lecture du film*, mais procède plutôt à une *désorientation par la lecture du texte*. Le spectateur «entre» dans l'univers diégétique déconcerté, indécis quant au degré de fictionalité du dit, et de ceux qui parlent. Il n'est pas étonnant que la Française puisse déclarer simultanément : «Je mens et je dis la vérité.»

### VII.5 SYNCHRONISATION, COREFERENCE

Au cours de mon analyse, j'ai pointé quelques phénomènes de coréférence qui témoignent de la minutie avec laquelle Alain Resnais, cet ex-monteur, conçoit la synchronisation des pistes visuelle et auditive. Il ne m'est pas possible de discuter ici l'intégralité des principes mis en œuvre pour établir une coréférence verbo-icnique dans les séquences en voix-over, car le film donne lieu à de multiples interactions. Dans un discours qui prend implicitement comme point de départ la supposée redondance de la plupart des voix-over au cinéma, Jean Mitry valorisait, dans son ouvrage paru six ans après la sortie du film de Resnais, la façon dont se nouent les rapports entre texte et image dans *Hiroshima mon amour* en constatant que «le texte n'est jamais l'équivalent verbal des images; il leur fait écho comme une sorte de correspondance intérieure, d'où il résulte que l'on dit *autre chose* en parlant des *mêmes choses*»<sup>809</sup>. Selon Mitry, le texte (mais il faudrait plutôt dire «la voix») se trouve du côté de l'intériorisation, même lorsque la locutrice se borne à décrire objectivement les faits. La richesse de cet apport mutuel doit beaucoup à la coréférence différée ou à l'intersection référentielle. Ce second processus explique pourquoi voix et image ne disent pas «les mêmes choses» tout en se recoupant.

Dans la première partie de l'évocation de Nevers où le texte introduit les images du *flash-back* sans être *over*, les informations visuelles complètent et précisent la signification des brèves phrases de la Française qui, ne s'étant pas encore abandonnée à la résurgence du passé, demeure laconique. Lors de cette phase où la coréférence s'établit de façon non simultanée, l'image donne à voir ce que le personnage n'ose dire, ou ce qu'il se contente d'évoquer de façon évasive. Par exemple, la Française répond au Japonais qui

809. J. Mitry, *Esthétique et psychologie du cinéma*, Vol. 2, *op. cit.*, p. 108.

l'interroge sur son amour de jeunesse: «Non, il n'était pas français.» Cette affirmation est suivie d'un plan où l'on voit, filmé depuis l'intérieur d'un café, un soldat que certains spectateurs (probablement plus en 1958 qu'aujourd'hui) identifieront, sur la base de son uniforme, comme appartenant au III<sup>e</sup> Reich. Il existe bien une zone d'intersection entre texte et image: la voix nous dit et l'image nous montre qu'il n'est pas Français. Toutefois l'image révèle combien la négation verbale est lourde de conséquences pour celle qui fut considérée comme une «traîtresse à la patrie»: ce «non-Français» est d'une part un Allemand, d'autre part un soldat.

Quant à l'annonce verbale du décès de l'amant, elle n'infléchit pas le contenu des images, mais s'accompagne de la monstration d'un endroit apparemment désert, un balcon baroque, dont l'apparition fait immédiatement suite à l'adjectif «mort» et s'imprègne, énigmatique, durant sept secondes sur la rétine du spectateur. L'association est posée entre ce lieu et le drame. Pour l'instant, aucune coréférence ne s'établit: la Française ne mentionne pas ce balcon qui, en soi, ne laisse rien présager. Si Resnais est, comme l'a formulé Benayoun, un «arpenteur de l'imaginaire», c'est à mon sens parce qu'il parcourt avant tout des espaces concrets qui, dans une représentation globalement subjective, deviennent des traces mnésiques: le balcon renvoie métonymiquement à la mort parce que, comme nous l'apprend la voix-over en coréférence simultanée sur cette même image (64 plans plus tard), «quelqu'un avait tiré d'un jardin». La relation métonymique est concrétisée par un brusque panoramique qui met en relation la berge sur laquelle gît le soldat agonisant et le balcon. La brèche qui existait entre ces deux espaces dans la construction de la diégèse est ainsi colmatée. Par conséquent, le retardement de la coréférence s'inscrit dans une entreprise de dramatisation. La participation active du spectateur qu'exige le montage audiovisuel de Resnais n'exhibe nullement l'aspect machinique du défilement des plans, mais produit véritablement l'émotion. La rétention du savoir et du sens contribue à mettre le spectateur dans la position désemparée du personnage qui explore l'abîme de sa mémoire.

En deçà des phénomènes de congruences sémantiques et référentielles, *Hiroshima mon amour* convoque une dimension rythmique qui, d'après les déclarations du cinéaste, lui importe beaucoup. Rares sont les films qui font à ce point primer la synchronisation sur le synchronisme. Celle-ci s'effectue aux deux



niveaux qui affectent profondément le discours filmique : la mise en cadre (travellings) et la mise en chaîne. Cette gestion de la temporalité n'est cependant pas dissociable d'une voix, celle de Duras lisant son texte, que Resnais pouvait réentendre à l'envi grâce à des bandes magnétiques – conditions d'écoute acousmatique qui n'ont peut-être pas été sans effet sur le caractère flottant des voix du film. L'unicité de la voix de l'écrivain correspond à l'homogénéité stylistique des répliques qui donne l'impression que la « voix narrative » traverse les voix de « tous » les personnages (en fin de compte, il n'y a pas vraiment d'autres personnages que la Française, le dialogue étant une sorte de prétexte au soliloque). Les paroles fixées offrent une durée, une scansion, une modulation qui constituent pour Resnais un étalon pour concevoir l'agencement des images, la durée des plans et la vitesse des déplacements effectués par le cadre et dans le cadre. Comme le note Laure Adler, « Resnais avait respecté le texte de Marguerite au point de calculer, chronomètre en main, le temps de ses travellings sur le « rythme *moderato cantabile* » de la phrase durassienne »<sup>810</sup>. Ce qui est valable pour le tournage l'est aussi pour la segmentation des plans, qui, dans un cinéma comme celui de Resnais, est envisagée dès le découpage. Dans une lettre du 7 août 1958 adressée à Marguerite Duras (voir le livret de l'édition DVD Arte), le cinéaste explique où il en est dans son travail : « Je viens d'attaquer le découpage. J'ai choisi le deuxième acte. C'est impressionnant d'entendre *voire voix me dicter les collures*. » Le rythme des images s'accorde à la segmentation syllabique du texte dit, comme l'illustre un bref extrait ostensiblement soumis à un réglage minuté de la synchronisation de mise en chaîne : il s'agit de cinq plans (101-105) du prologue montrant de manière toujours plus rapprochée et sous différents angles le Palais de l'Industrie, seul bâtiment à être resté debout après le bombardement. À chacune de ces images qui durent environ une seconde (sauf la dernière, qui fait office de point d'orgue avant le retour à l'image des corps enlacés) correspond exactement un terme du texte *over*. Ce dernier se voit dès lors fragmenté comme suit dans sa relation à l'image : « Pourquoi/ nier/ l'évidente/ nécessité/ de la mémoire ? » On voit bien comment un mot peut, chez Resnais, « dicter une collure ». Ce n'est pas un hasard si le montage est mis en évidence à ce moment-là du film : l'exhibition

810. L. Adler, *op. cit.*, p. 525.

du travail de concaténation des images participe de la signification même de la phrase, la fragmentation du flux visuel étant associée au travail de la « mémoire », une notion qui, littéralement, « a le dernier mot » dans cette séquence. La progression résultant des réductions successives de l'échelle des plans (notamment grâce à un raccord dans l'axe entre l'avant-dernier plan et le dernier) semble signifier une percée vers l'intériorité. À nouveau, Resnais fait reposer la subjectivisation sur la monstration d'un espace concret, d'une vue documentaire. L'ultime plan montre la pointe de la coupole du bâtiment, dont il ne subsiste plus que l'armature de fer, filmée en contre-plongée. Resnais nous fait ainsi pénétrer au cœur de la convergence des multiples fils métalliques qui s'enchevêtrent de façon arachnéenne. Cette image semble figurer les réseaux tissés par le cerveau, qui intéresse surtout Resnais du point de vue de la « mémoire », mot prononcé simultanément à la visualisation de l'objet.

#### VII.6 L'IMMERSION MALGRE TOUT

La facture très construite du montage, le réglage précis des phénomènes de synchronisation et de coréférence ainsi que l'artificialité de la diction d'Emmanuelle Riva pourraient laisser croire que le film de Resnais procède à une froide mise à distance de l'audio-spectateur, exhibant la mécanique cinématographique en évacuant toute émotion. Certains commentateurs analysent l'usage de la voix dans *Hiroshima* et *Marienbad* comme un « effet de distanciation » au sens brechtien, c'est-à-dire comme un élément destiné à faire obstacle à l'immersion diégétique du spectateur et à l'identification au personnage<sup>811</sup>. Or j'ai souligné à plusieurs reprises combien l'affirmation de l'humain qui s'effectue à travers l'expression vocale occulte la dimension machinique de la représentation, favorisant le rapport direct entre le spectateur et le monde du film. Il serait à mon sens plus juste de parler de distanciation à propos d'un film comme *On connaît la chanson*, où la déliaison vocale implique une exhibition de la matérialité phonographique. Dans le prologue d'*Hiroshima mon amour*, le potentiel de la vocalité introduit des indices de subjectivité décelables par le spectateur dès

811. Voir par exemple R. Blaetz, « L'impiego della retorica in due film di Alain Resnais », *Cinema Nuovo*, août-septembre 1982, pp. 54-55.

lors que celui-ci accepte le «pacte» initial d'un discours défini par l'emphase lyrique. Keith Cohen, qui traite du «plaisir sonore» dans les films de Resnais en évoquant la «possibilité d'une régression narcissique à une expérience prénatale de l'union avec la mère» (d'autant plus forte ici que c'est une voix féminine qui domine l'énonciation verbale *over*), montre que ce plaisir est amplifié par l'utilisation enveloppante de la voix-over des monologues, l'éviction du corps de la locutrice pouvant être interprétée comme l'un des mécanismes soustractifs qui, pour Freud, composent le principe de plaisir<sup>812</sup>.

Le prologue d'*Hiroshima mon amour* n'a donc pas besoin de capitaliser sur l'établissement d'une figure fictionnelle pour conférer à la voix-over sa charge émotive. Marie-Claire Ropars dit de la voix dans le film de Resnais que «l'artifice s'accepte parce que la voix se trouve renvoyée à un discours intérieur dont elle serait l'émanation transposée»<sup>813</sup>. Ropars comprend la notion de «discours intérieur» dans le sens d'un monologue effectué au sein de l'univers diégétique. On a toutefois vu que, si l'on prend ce terme dans le sens de Boris Eikhenbaum, la «transposition» extériorisée s'origine dans l'esprit du spectateur, non dans la «psychologie» d'un personnage du film. Cette compréhension du «discours intérieur» me paraît fondamentale pour comprendre pourquoi, en dépit de la non-visualisation du locuteur, la «projection» du spectateur dans le monde du film est possible, et se voit même renforcée en raison de la désincarnation du prologue et de l'importance de la dimension rythmique. La «musicalité» des films de Resnais se manifeste tant au niveau de la récurrence de motifs visuels, dans le choix des musiques que dans le choix de textes qui recourent de façon systématique à des procédés de répétition (le «film industriel» *Le Chant du Styryène* est par exemple accompagné d'alexandrins *over* écrits par Raymond Queneau). La juxtaposition des travellings, dont on a vu que la vitesse était calquée sur le débit d'une voix, participe également de la gestion rythmique. Les fondus enchaînés qui relient ces travellings renforcent la fluidité du perçu. Ce «signe de ponctuation» n'est pas seulement un élément de démarcation, il est aussi le lieu de la superposition des images,

812. Keith Cohen, «Pleasures of Voicing: Oral Intermittences in Two Films by Alain Resnais», *L'Esprit créateur*, Vol. 30, N° 2, 1990, p. 58.

813. Marie-Claire Ropars, *De la littérature au cinéma. Genèse d'une écriture*, Paris: Armand Colin, 1970, p. 150.

d'un flottement qui s'accorde à la présence-absence de la voix-over. Les travellings avant provoquent par ailleurs une sollicitation quasi corporelle du spectateur, ainsi que l'a montré Laurent Jullier, qui compte ce procédé parmi les « figures de l'immersion » du cinéma<sup>814</sup>. Toutes ces caractéristiques stylistiques participent d'une exploitation des vertus hypnogènes de la musicalité des sons en général, vertus dont Laurent Guido a montré qu'elles constituaient l'un des cadres de réflexion sur la musique au cinéma dans les années 20, notamment sur la base de l'hypothèse formulée par Bergson dans son essai de 1888 à propos de l'action de la musique :

« Le rythme et la mesure suspendent la circulation normale de nos sentiments et de nos idées en faisant osciller notre attention entre des points fixes et s'emparent de nous avec une telle force que l'imitation, même infiniment discrète, d'une voix qui gémit suffit à nous emplir d'une tristesse extrême. »<sup>815</sup>

En emportant le spectateur dans le mouvement général insufflé par diverses organisations rythmiques superposées (le montage, les mouvements d'appareil, le phrasé du texte et les motifs musicaux), *Hiroshima mon amour* démultiplie l'impact émotif des discrètes fluctuations vocales d'Emmanuelle Riva. Alain Resnais use volontairement du pouvoir hypnotique des sons auxquels le spectateur est totalement livré, plus encore qu'à l'image. Il déclare par exemple, à propos d'*Hiroshima mon amour* : « Je tenais à composer une espèce de poème où les images ne serviraient que de contrepoint au texte. Car, pour reprendre la phrase d'André-S. Labarthe : « Au cinéma, je peux fermer les yeux, mais je ne peux pas me boucher les oreilles ! »<sup>816</sup> Entendre une voix-over, c'est en quelque sorte regarder des images les yeux fermés, comme si elles venaient du dedans, se confondaient, suscitées par le référent verbal, avec le « discours intérieur » du spectateur.

Alain Resnais formule d'ailleurs explicitement ses intentions quant à l'utilisation de la voix-over : il s'agissait pour lui d'« arriver à créer chez le spectateur une sorte d'hypnose »<sup>817</sup>. Cette démarche

814. L. Jullier, *L'écran post-moderne...*, op. cit., pp. 74-75.

815. Henri Bergson, *Essai sur les données immédiates de la conscience*, cité par L. Guido, op. cit., p. 478.

816. Cité par B. Pingaud, *Premier Plan*, N° 18, 1961, p. 40.

817. Cité par M. Firk, « *Hiroshima mon amour*, film scandaleux? », *Les Lettres françaises*, N° 773, 1959.

correspond justement à ce qui, pour Bertolt Brecht, caractérise les «effets anciens de distanciation» basés sur la «suggestion hypnotique», qu'il distingue de ceux qu'il tente lui-même de mettre en œuvre<sup>818</sup>. D'une part, nous avons vu que le film de Resnais n'enrayait pas l'identification, mais se voyait traversé par la constitution progressive de la psyché du personnage féminin. D'autre part, l'hypnose échappe à l'objectif brechtien consistant à laisser au spectateur la plus grande liberté interprétative possible: au lieu de mener à une *prise de conscience*, le film de Resnais soumet plutôt son audiospectateur à l'*emprise d'une conscience*. La prose incantatoire de Duras récitée par Riva s'oppose à une diction qui, selon Brecht, «doit se débarrasser du ton psalmodiant et des rythmes fixes qui bercent le spectateur»<sup>819</sup>. On peut donc dire que, paradoxalement, l'éviction partielle de la représentation visuelle des personnages et leur mise à distance s'accompagnent dans *Hiroshima mon amour* d'une adhésion émotive de l'audiospectateur. Même désincarnée, la voix dégage une aura de subjectivité sur l'ensemble du film, s'assimile à la «voix narrative» qui porte le film et nous emporte. Si Resnais peut exp(l)oser les corps de ses personnages comme il le fait dans le prologue, les réifiant par la fragmentation du cadre qui les transforme en un lacs de lignes géométriques, c'est parce que les affects se déploient hors de ceux-ci, au travers d'«attitudes propositionnelles» globalement définies par le texte *over*. À propos de la dimension émotive des films de Resnais, Gilles Deleuze distingue la «psychologie des sentiments purs» qui y est à l'œuvre de la psychologie traditionnelle des personnages au cinéma<sup>820</sup>. Il cite dans ce contexte une explication où René Prédal dément la thèse d'un *Verfremdungseffekt* resnaisien :

«Si Brecht obtenait au théâtre ce résultat par la distanciation, Resnais provoque au contraire une véritable fascination. Cette sorte d'hypnose [...] dirige l'attention du public sur les seuls sentiments qui animent le héros.»<sup>821</sup>

Peut-être ne faut-il pas opposer trop strictement distanciation et fascination, car la validité de cette dichotomie dépend de la

818. Bertolt Brecht, *L'Achat du cuivre. Petit organon*, Paris: L'Arche, 1999 [1964], p. 206.

819. *Ibid.*, p. 208.

820. G. Deleuze, *Cinéma II. L'Image-temps*, op. cit., p. 163.

821. R. Prédal, *Études cinématographiques*, N° 64-68, op. cit., p. 163.

conception que l'on se fait de la notion de Brecht. Resnais a lui-même souligné la polysémie de cette notion en donnant une acception du terme qui est révélatrice de la fonction conférée à la voix-over dans *Hiroshima mon amour*:

«Il faudrait s'entendre sur une définition de la distanciation. C'est un mot qui recouvre tant de choses. Si vous êtes happé, c'est grâce à la distanciation, c'est parce qu'on demande une participation de votre part.»<sup>822</sup>

Il est néanmoins évident que le travail esthétique de Resnais ne correspond pas totalement aux objectifs brechtiens, le dramaturge allemand accordant à la voix une fonction commentative déagée de la diégèse. On saisit par contre combien, chez Resnais, l'exigence renforcée d'une «coopération textuelle» de l'audiospectateur induite par les référents verbaux et les différentes relations qu'ils établissent avec l'image (par retardement, contraste ou intersection partielle) participe de l'immersion diégétique du spectateur. Même si l'on oscille entre la méga-énonciation fictive et la constitution d'un personnage présenté comme étant à la source du discours audiovisuel, l'adhésion fonctionne, puisqu'on demeure dans le pôle humain de l'énonciation filmique.

#### VII.7 LA VOIX (RE)INCARNEE: LE «REMAKE» DE NOBUHIRO SUWA

*H Story*, sorte de «*remake*» du film d'Alain Resnais réalisé par le cinéaste japonais Nobuhiro Suwa en 2001, témoigne de la fascination qu'*Hiroshima mon amour* continue d'exercer aujourd'hui. Il me paraît intéressant de terminer sur ce prolongement récent de l'œuvre de Duras et Resnais dans la mesure où le troisième long métrage de Suwa se distingue radicalement de son modèle du point de vue de l'usage qu'il fait de la voix, éclairant ainsi par la négative les effets provoqués par l'esthétique resnaisienne. En effet, le film de Suwa, histoire d'une tentative avortée de reconstitution maniaque d'*Hiroshima mon amour* à l'époque contemporaine et en couleurs, écarte systématiquement la conjonction d'une voix-over et de la visualisation du passé: les répliques de la Française, interprétée ici par Béatrice Dalle, sont prononcées en voix synchrone

822. Cité par J.-L. Leutrat, *Hiroshima mon amour*, p. 73.

ou, plus rarement, sur des plans noirs. Nevers n'existe donc qu'à travers le référent verbal, le passé du personnage étant définitivement hors-champ. Ce qui importe dans le film de Suwa, ce n'est pas tant l'histoire de Duras que la façon dont on peut raconter oralement cette histoire, c'est-à-dire la tentative d'*incarner*, aujourd'hui, le personnage qui déclame son texte. La caméra est donc continuellement braquée, en plans fixes, sur le visage et les gestes de Béatrice Dalle, montrant avec insistance comment celle-ci échoue à s'approprier un texte qui, de toute évidence, n'est pas fait pour être dit continuellement *in*. Alors que chez Resnais les corps s'effacent pour laisser la place à un univers mental, *H Story* se concentre sur l'expression corporelle de la vocalité. La démarche du cinéaste japonais est donc inverse à celle qui est mise en œuvre dans *Hiroshima mon amour*, où le texte a une existence presque parallèle aux êtres qui en sont les vecteurs. C'est pourquoi l'élocution de Dalle n'a rien de la diction claire, tantôt martelée, tantôt flottante, d'Emmanuelle Riva, mais relève plutôt d'un chuchotement peu articulé et parasité par le contexte concret de la situation d'énonciation visualisée (mauvaise acoustique des scènes d'extérieurs, bruits environnants). Les paroles sont à la limite de l'intelligibilité, exigeant parfois que l'audiospectateur se remémore le texte de Duras pour être comprises. L'interlocuteur japonais parlant ici français, le film a pu être tourné en prise de son directe, ce qui contraste fortement, au niveau de la matérialité tant vocale que phonographique, avec l'enregistrement postsynchronisé de Resnais. Même dans les rares passages en voix-over de *H Story*, la performance vocale résulte ostensiblement d'un enregistrement (on entend généralement la caméra tourner à l'arrière-plan sonore, même s'il n'y a pas d'image), et d'une improvisation de l'actrice.

Alors que, dans *Hiroshima mon amour*, nous ne voyons jamais la Française jouer dans la production internationale à laquelle elle participe et qui motive sa présence dans la cité nippone, Suwa prend le parti de considérer l'ensemble de son film comme un tournage. Le référent principal des paroles et des images de Suwa n'est pas l'aventure de Nevers, mais le film de Resnais, qui apparaît sous la forme d'images fixes, à l'instar de certaines photographies de victimes irradiées dont il est lui-même composé. Le texte de Duras est discuté en tant que tel, présent dans sa dimension matérielle d'objet (Béatrice Dalle sort parfois le scénario de sous la table lorsqu'elle joue la scène du tea-room) ou sous forme écrite (mentions

sur fond noir). Le texte est considéré comme quelque chose de *fixé* dont la réactualisation dans l'*oralité* provoque une labilité qui est d'autant plus perceptible que l'audiospectateur peut comparer les voix entendues aux souvenirs qu'il a des voix du film de Resnais.

À une époque qui connaît une prolifération des films à voix-over et enchâssements et une banalisation des « jeux sur le récit » tels qu'en présentaient de façon novatrice les films d'Alain Resnais durant la période 1960-1980 (avec une résurgence momentanée de ce type de problématique dans *Smoking! No Smoking*), Nobuhiro Suwa préfère envisager le texte de Duras du point de vue de son incarnation vocale. Il suit ainsi l'évolution d'une certaine « modernité » cinématographique qui, tout en s'inscrivant dans l'héritage resnaisien de l'appropriation d'un texte littéraire, met en évidence l'acte même de profération de la parole plutôt qu'elle ne tend vers l'acousmatisation des voix. Ainsi la trilogie réalisée par Jean-Marie Straub et Danièle Huillet (1998-2003) à partir des textes de Vittorini, romancier dont l'écriture présente une musicalité que Resnais rapprochait de celle du style durassien, repose-t-elle sur un principe d'ancrage de l'émission vocale dans une présence physique de l'« acteur » (qui déclame plus qu'il n'« agit »). Cette démarche ne vise pas à asservir le réalisme audiovisuel à la psychologisation des personnages, mais à montrer comment la vocalité et l'inscription dans une situation d'énonciation donnée « travaillent » le texte, celui-ci conservant son autonomie par rapport à l'univers diégétique. C'est également ce qu'entreprend Suwa, mais de façon moins radicale puisqu'il motive diégétiquement (grâce à l'argument du « film dans le film ») les situations d'énonciation du texte de Duras.

On observe une incidence de la réflexivité de *H Story* sur les voix qui s'y font entendre. D'une part, le cinéaste est lui-même présent dans son film, et s'entretient avec l'actrice par l'intermédiaire d'une traductrice. Cette voix « auctoriale » se situe toutefois aux antipodes des voix-over de Cocteau, Guitry, Truffaut, etc. qui dominent l'organisation des films de ces cinéastes. Ici, Suwa n'est qu'un personnage parmi d'autres, et les difficultés qu'il rencontre pour se faire comprendre de Béatrice Dalle témoignent au contraire de son impuissance par rapport au déroulement de son « *work in progress* ». Contrairement à *Hiroshima mon amour*, où la voix d'Emmanuelle Riva semblait générer les images et guider le récit, le film de Suwa ne comporte aucune personnalisation vocale de ce type. En raison de l'absence d'instauration d'un méga-énonciateur fictionnel, le



niveau de la méga-énonciation fictive est plus fortement perçu dans ses rapports à son origine mécanique. Or Suwa laisse dans son film quantité de traces qui renvoient manifestement aux conditions machiniques du tournage. *H Story* présente en effet une facture brute qui s'apparente à celle de *rushes*, les plans étant séparés par d'importantes sautes, soulignées par l'absence d'étalonnage. Cette exhibition de la technique au niveau du marquage matériel de l'énonciation filmique concerne également les sons : marquage prophonographique (les bruits de l'équipe située à proximité du perchiste), phonographique (les claps sont effectués au début des prises alors que retentit le bip destiné à la synchronisation, certaines phrases sont inaudibles parce que l'actrice ne les profère pas dans l'axe du micro) et postphonographique lorsque des scènes dialoguées ne sont subitement plus sonorisées. De telles déliaisons par suppression de la voix dénaturalisent les effets de la synchronisation vocolabiale en rappelant qu'elle résulte *in fine* d'un travail effectué à l'étape du montage.

Ainsi le « *remake* » délibérément avorté d'*Hiroshima mon amour* se distingue-t-il radicalement de l'original en présentant une énonciation filmique ouvertement discursive, alors que le film de Resnais humanise la représentation en soumettant le flux des images au phrasé d'un texte porté par la voix d'un personnage. *H Story* ne dégage rien d'hypnotique : la voix, enchaînée au *hic et nunc* d'une énonciation entravée par l'inaptitude de l'actrice, casse toute continuité rythmique. Si Alain Resnais, en dépit de la méticulosité dont il témoigne au niveau du montage, fait de son film l'expression fictionnelle d'une expérience humaine, Nobuhiro Suwa écrase la manifestation orale du texte de Duras sous l'appareillage nécessité par la représentation cinématographique. Le rapport de l'humain à la machine s'inverse, transformant fondamentalement le régime énonciatif du film et les modalités des manifestations vocales.



## ÉPILOGUE·S

La réflexion théorique et historique menée tout au long de cet ouvrage sur les différents degrés d'appartenance des voix du cinéma au régime de l'*oralité* ou à celui l'*écrit* a permis d'examiner dans une optique «sémio-pragmatique» à la fois la représentation (le film comme «texte» dans son rapport à son destinataire *via* les inférences qu'il suscite) et les différents paramètres qui définissent les «dispositifs» du cinéma parlé et parlant (locuteur vivant, place de la machinerie, position du spectateur, etc.). À travers une comparaison entre le «cinéma» bonimenté et la voix-over du parlant, on a pu constater certaines parentés et dissemblances dont on a dégagé des critères d'analyse tels que le type et le degré d'intégration, la valeur (pseudo)attractivité, la fonction dominante du langage, le rapport à la diégèse visualisée, le mode de synchronisation et de coréférence, les niveaux de marquage énonciatif, etc. En ce qui concerne le pré-cinéma, il subsiste tout un ensemble de pratiques à examiner du point de vue de l'illusion du synchronisme audiovisuel, que je n'ai abordé que très ponctuellement: il s'agit des «machines parlantes» au sens large, expression qui comprend autant les automates que les tentatives de couplage du phonographe avec l'image. L'usage de ces techniques a en effet contribué à créer un fondement anthropologique commun à différents spectacles qui se perpétue dans le cinéma parlant dominé par la voix synchrone. D'ailleurs, l'émergence des «nouvelles technologies» n'a pas fondamentalement modifié la place et la fonction attribuées à la voix dans le domaine de l'exploitation cinématographique, ni même au sein de l'institution du film fictionnel puisque, même dans les productions recourant massivement ou intégralement à l'image de synthèse, la voix continue d'être exploitée comme la trace d'une expression humaine individuelle, reconnaissable dans sa

nature prophanographique. C'est pourquoi la typologie des régimes vocaux que j'ai proposée conserve sa validité dans la grande majorité des discours et des pratiques actuelles.

On distingue premièrement la voix-action, régie par le principe du synchronisme qui assure un simulacre réaliste de l'humain en occultant toute part machinique; deuxièmement, la voix-attraction, issue des spectacles vivants et susceptible d'être mise en scène sur la base d'une interactivité fictive dans un film parlant; enfin la voix-narration (ou «voix-explication» en fonction du mode d'organisation textuelle choisi), où le verbal se met au service d'une intégration au discours filmique pour donner l'impression qu'il est à l'origine de ce dernier. Globalement traversée par la bipolarité attraction *vs.* narration, mon approche de la voix renoue avec l'étude du cinéma des premiers temps qui s'est développée dans le sillon «post-Brighton», où la théorie du «cinéma des attractions» permit de faire converger des préoccupations historiques, esthétiques et narratologiques. Sur cette base, j'ai pu avancer des critères pour l'analyse filmique et pour une histoire (et une historiographie) de la voix au cinéma. En axant prioritairement mon étude sur les conséquences théoriques de l'analyse historique des faits et des discours, je me suis borné à livrer quelques repères ponctuels de cette «histoire de la voix» qui demeure encore très largement à faire, mais pour la constitution de laquelle j'espère avoir ouvert quelques pistes.

Les modalités d'utilisation du verbal dont je rends compte dans ma typologie des régimes vocaux n'ont pas attendu la constitution du «langage cinématographique» pour s'adjoindre à la projection d'images fixes ou animées. La voix-action est certes difficilement envisageable à l'époque du «cinéma» parlé, si ce n'est sous la forme, dont l'existence est attestée au début du XX<sup>e</sup> siècle, du doublage des personnages de l'écran. La représentation dominante à laquelle la voix-action donne lieu, celle de l'unité vocale et corporelle d'un locuteur, renoue en revanche avec la tradition des «machines parlantes». Un visage isolé de son environnement, une bouche qui remue et, simultanément, une voix synchronisée avec les mouvements labiaux selon un degré de précision que le public juge convenable en regard de certaines normes qu'il a assimilées: tels sont les aspects concrets d'une représentation qui correspond à l'archétype d'un fantasme techniciste, scientifique mais aussi «poétique» (au sens large) qui obsède depuis des siècles l'esprit

humain<sup>823</sup>. La voix synchrone du cinéma n'est que l'une des réalisations de cette aspiration prométhéenne, qui, en fonction des différents usages de la déliaison vocale, se voit soit mise à mal (en cas de marquage de la matérialité phonographique), soit exacerbée dans sa dimension humanisante (par exemple lorsqu'une voix-over est attribuée à un personnage focal). De nombreux cas intermédiaires de voix déliées restent à explorer, notamment sous l'angle de l'effet d'«inquiétante étrangeté» dont il a été question à quelques reprises dans cet ouvrage. Il est par exemple fort probable que l'angoisse suscitée par la déliaison fondamentale, indice d'une perte de l'intégrité «corporelle» du locuteur auquel s'identifie l'audio-spectateur – ce dernier étant dès lors confronté à la mécanicité du dispositif, ainsi que l'envisagent indirectement les commentateurs qui ont abordé la question du bruit du projecteur, soit généralement reversée sur les caractéristiques de l'univers diégétique par ce que j'ai proposé d'appeler la «mise en phase». À cet égard, le genre fantastique semble être le lieu privilégié de la voix déliée.

La voix-attraction a été définie sur la base de trois critères qui s'appliquent pleinement aux spectacles vivants dont dérivent certains films parlants: elle est labile, car non réitérable sous une forme strictement identique en raison de sa dépendance envers les circonstants éphémères de la performance analysés par Zumthor; l'usage de la fonction référentielle du langage n'y a pas pour but de créer un univers diégétique, mais d'exhiber les composantes du spectacle; elle est ancrée dans le présent de son énonciation et n'a de sens que par rapport à lui. Dans le cinéma parlant (ou à la radio, à la télévision), certains usages vocaux sont, en dépit de leur nature de sons fixés, indéniablement plus attractionnels que d'autres (même s'il ne s'agit là que d'une «pseudo-attraction») dans la mesure où ils simulent l'interactivité ou la modularité.

### VOIX-NARRATION ET ENONCIATION

La voix-narration participe de l'organisation discursive de la représentation audiovisuelle. Ainsi, les techniques de fixation lui

823. Voir A. Boillat, «*L'Ève future...*», *art. cit.*, ainsi que le chapitre 2 de *Les Voix du cinéma. Boniment, voix synchrone, déliaison et voix-over*, où je traite autant de discours savants portant sur la fabrication d'automates ou sur l'utilisation du Phonoscope que d'œuvres littéraires qui exploitent sur le plan narratif cet imaginaire (*L'Ève future, Le Château des Carpathes, L'invention de Morel*).

sont favorables, puisqu'elles permettent d'affiner et de pérenniser le réglage des processus rythmiques (synchronisation) et sémantiques (coréférence) pour l'analyse desquels j'ai développé certains outils. Aussi ai-je préconisé de concevoir les relations entre le verbe et l'image comme un phénomène de réciprocité (récusant ainsi la pertinence de la notion de « redondance »), de distinguer la signification de la référence et de croiser les critères de l'analyse sémantique avec les principes de la synchronisation.

Par ailleurs, la présence-absence inhérente à la technique phonographique s'accorde à la disparition du corps du sujet parlant *over*. L'origine « personnalisée » de la voix-over est alors assimilée par l'audiospectateur à une instance discursive qui générerait l'ensemble du film. En se substituant au « discours intérieur » tel que l'envisageait Boris Eikhenbaum, la voix-over actualise en quelque sorte par extériorisation un discours verbal sous-jacent à l'organisation visuelle, se confondant ainsi fictivement avec le siège des activités mentales de l'audiospectateur. Ces processus ne s'instaurent toutefois que si la voix-over ne présente pas, en raison d'une déliaison (sous forme de dissociation, de plaquage, etc.) ou d'altérations (post)phonographiques, une certaine « rugosité » qui provoque une forme d'étrangeté, et donc d'altérité. Toutefois, la voix-over verbalise généralement dans l'oralité ce que l'audiospectateur est incité à considérer, au vu des circonstances perceptives de la réception et du caractère enveloppant de l'univers sonore, comme l'expression du travail cognitif qu'il est lui-même en train d'effectuer lors de la lecture de l'image. L'importance de la voix-over comme facteur d'immersion diégétique est ainsi indissociable des questions d'énonciation filmique, invariablement soulevées depuis Albert Laffay lorsqu'il s'agit d'analyser les effets d'une présence *over*. Tirailé entre l'humain et la machine, le modèle dont j'ai fait l'hypothèse se veut bifide : d'un côté il se démarque radicalement des conceptions humanisantes jusqu'ici dominantes dans le champ des études cinématographiques (chez Odin, Jost, Châteauvert), de l'autre il réintègre lesdites conceptions en les considérant comme des processus secondaires relevant de la fictivité. Ainsi, la façon dont la voix-over a été traditionnellement théorisée fait elle-même office de « fait de réception » à partir duquel on peut tenter d'établir certaines catégories de processus inférentiels.

Un examen de la notion de « voix narrative » a permis de relever des similitudes entre le lecteur d'un roman et le spectateur d'un film au niveau de la perception d'une origine du discours. Or,

lorsque les théoriciens du cinéma empruntent le modèle de l'énonciation narrative appliqué au texte scriptural, ils sous-estiment grandement la part machinique de la production de la représentation audiovisuelle, y compris en ce qui concerne les voix du film. Les notions de «voix narrative» ou d'«énonciateur filmique» me paraissent pourtant devoir être étudiées en ce qu'elles sont représentatives des conditions de réception d'un récit (scriptural ou filmique). Si le spectateur a l'impression qu'«on lui parle» au cinéma, même lorsque le film est muet, on est en droit de se demander ce qu'il en est lorsqu'il entend une voix-over qui, de façon plus ou moins explicite, s'adresse à lui. Le statut *over* implique une dissociation du corps et de la voix qui se fait l'écho de la situation perceptive du spectateur dans le dispositif. C'est sur cette base que s'effectue le lien pragmatique, soit par reversement sur l'univers diégétique des effets de la déliaison, soit par l'usage de la voix-over dont les propriétés sont tout à fait adaptées aux conditions fondamentalement acousmatiques de la réception du film.

La conception de l'énonciation cinématographique que je prône, fortement redevable du contexte général posé par Metz dans *L'Énonciation impersonnelle*, est dérivée de la conjonction des deux catégories à l'aune desquelles j'ai appréhendé le rôle de la voix au cinéma : les conditions matérielles de la production mécanique et l'élaboration fictive d'un pôle humanisant. Sur cette base, j'ai tenté de renouveler l'approche énonciative en mettant l'accent sur les dimensions machinique et pragmatique. Le modèle que j'ai établi démontre que l'opposition de l'humain et de la machine constitue un axe d'analyse productif du point de vue des phénomènes vocaux, tant au niveau du dispositif cinématographique, de la représentation audiovisuelle que du discours filmique. En fait, il permet de prendre à la lettre l'expression de Christian Metz selon laquelle il y aurait un «discours sous l'histoire», celui-ci se «déguisant» en celle-là. On peut en effet comprendre le terme «discours» au sens des conditions matérielles de production de la représentation, et celui d'«histoire» comme un phénomène de surface, une «illusion» – que j'ai tenté de cerner à travers l'opposition fictivité *vs.* fictionalité – créée par l'organisation discursive du film afin d'humaniser sa source énonciative. De cet effet sur le spectateur découle la spécificité de la voix : qu'elle émane du bonimenteur ou d'un narrateur *over*, la parole a pour fonction, en plus de constituer un apport sémantique

important, d'affirmer la présence «vocale» de l'humain, niant ainsi la mécanicité de la production audiovisuelle.

L'importance des conséquences «sémio-pragmatiques» de la nature phonographique des voix du cinéma parlant m'est apparue par contraste avec la performance bonimentorielles. La voix émise par les haut-parleurs dans la salle est le résultat d'une chaîne d'opérations de sélection/modification des sons dont il est nécessaire de tenir compte dans l'analyse, comme il est indispensable de partir du principe que le spectateur, lui aussi, peut être amené à en tenir compte. Le «savoir de l'arché» relatif à la composante sonore s'exerce sur deux facteurs qui font intervenir des connaissances portant sur la nature *représentationnelle* des sons. L'un contribue à humaniser toute présence vocale (c'est le savoir lié à la nature indicielle de l'enregistrement sonore, trace d'un événement phonographique effectif), l'autre permet une prise de conscience du caractère machinique des sons. Cette dernière est quasi généralisée aujourd'hui, où les spectateurs de cinéma sont pour la plupart familiarisés avec la manipulation sonore, dont les applications s'étendent du plus simple maniement des potentiomètres d'un amplificateur sur une chaîne stéréo aux tables de mixage virtuelles que proposent des programmes informatiques fort répandus.

### LE MONDE DU FILM

Les fortes potentialités de coréférence verbo-iconique dont dispose la voix-over la destinent majoritairement à la fonction référentielle du langage. C'est pourquoi j'ai voulu traiter de la parole *over* comme d'un acte de référence qui se combine avec celui de l'image pour produire du sens. Sur cette question des relations verbo-iconiques, on constate un relatif désert théorique que ne laisse pas présager le discours récurrent sur la prédilection dont témoigneraient les études sur le son au cinéma pour le langage verbal. Je me suis contenté, dans cet ouvrage, de proposer une typologie des phénomènes de coréférence verbo-iconique et de discuter ponctuellement certains aspects sémiologiques résultant de la coprésence (simultanée ou différée) du mot et de l'image. Il s'agirait là aussi de pousser plus avant l'élaboration théorique, notamment en repartant des considérations sémiologiques émises dans les années 70 par Roland Barthes, Christian Metz ou Umberto Eco.



Par contre, j'ai mis l'accent sur la participation du verbal à la construction de l'univers diégétique, ce qui m'a conduit, au vu du statut singulier du locuteur *over* dont la voix peut relier des «mondes possibles» distincts, à rediscuter le qualificatif «extradiégétique» (dont on a vu que la pertinence pour notre objet était bien inférieure à celle du couple homodiégétique/hétérodiégétique) et à proposer une conception pluridiégétique qui me paraît plus adéquate pour rendre compte de certains enchâssements narratifs. Ce type d'approche s'avère aujourd'hui nécessaire pour analyser des films qui mettent en scène des univers virtuels, c'est-à-dire qui prennent acte de la multiplicité des mondes à laquelle le consommateur de productions audiovisuelles actuel est confronté. Certes, dans le montage parallèle griffithien déjà, le téléphone met en relation des espaces diégétiques disjoints. Néanmoins, les voix qu'entendent dans leur combiné ou par transmission radio les protagonistes de *Matrix* ne peuvent pas être uniquement analysées en termes de «voix acousmatiques» ou de «point d'écoute», car elles proviennent d'un autre monde, inclus dans la mégadiégèse mais distinct de celui qu'habitent les auditeurs diégétiques. Comme les chants dans *Avalon*, ces voix constituent la condition d'accessibilité à un monde régi par d'autres lois. Médiatisée par une technologie représentée de façon partiellement réflexive dans le film, la voix trouve son potentiel de communication ravivé: elle relie les êtres et les mondes.

Cette réflexion s'est accompagnée d'une mise au point terminologique et théorique concernant la définition de la notion de «diégèse», nécessaire pour dissiper certaines confusions qui remontent à l'ouvrage *Figures III* de Genette. Les précisions apportées ici poursuivent en quelque sorte le travail entrepris dans une optique différente (mais néanmoins parente) par André Gaudreault dans *Du littéraire au filmique* à propos du couple *diègèsis/mimèsis*: alors que le théoricien québécois jugeait utile d'en passer à nouveau par les textes de Platon et Aristote pour discuter avec plus de rigueur une typologie galvaudée à force d'être reprise, j'ai quant à moi tenu à revenir sur la conception du filmologue Étienne Souriau, dont Genette s'inspire partiellement (et quasi tacitement). En retournant au domaine des études cinématographiques par-delà l'analyse littéraire, on constate à quel point les films constituent des objets complexes du point de vue de la logique modale, et cela notamment en raison de la richesse des matières de l'expression du médium et des libertés qu'offre le procédé de la voix-over.

L'importance de la voix-over en termes de «diégétisation» est manifeste dans la mesure où elle intervient très majoritairement dans la phase liminaire des films, endossant ainsi une fonction centrale d'instauration de l'univers. C'est pourquoi, prolongeant les travaux de Gardies ou d'Odin sur les débuts de films, j'ai mis l'accent en plusieurs occasions sur les interventions initiales de la voix, *live* puis *in* dans les prologues de type scénique, *over* dans *Le Roman d'un tricheur* ou dans *Hiroshima, mon amour*. Cet usage inaugural n'a pas fléchi dans les productions standardisées récentes comme *Le Seigneur des anneaux: La Communauté de l'anneau* (Peter Jackson, 2001), où la voix-over nous immerge dès le générique dans un passé des origines («Tout commença avec la forge des Grands Anneaux...», «L'histoire devint légende, la légende devint mythe»), *Spider-Man* (Sam Raimi, 2002), *Troie* (*Troy*, Wolfgang Petersen, 2004), *La Guerre des mondes* (*War of the Worlds*, Steven Spielberg, 2005) ou encore le film d'animation *Cars* (John Lasseter, 2006). Ces manifestations over momentanées à l'entame du film révèlent la fonction essentiellement contractuelle de la voix et les connotations «originelles» qui lui sont appliquées.

«Médium» des premiers conteurs, la voix apparaît comme l'instance première, quasi divine, qui donne accès au monde de la fiction. Nous avons vu, en recourant aux écrits de divers «poéticiens» (Weinrich, Hamburger, Ricœur) dont l'observation s'est avérée fort utile pour comprendre le fonctionnement de la voix-over, combien l'appréhension des temps passés peut être différenciée en fonction du statut fictionnel ou non fictionnel des films: les fictions renvoient à une époque immémoriale qui rejoint l'inexistant et postule une réception sur le mode du «comme si»; la voix des documentaires, dont l'énonciation tente en général de s'ancrer dans le présent, use de la valeur prétéritive de ses verbes pour authentifier le montré. Dans l'un et l'autre cas, la voix-over constitue un agent intégrateur: toute l'organisation discursive semble portée par le verbal.

Qu'il s'agisse de l'élaboration diégétique, de la construction d'une figure d'«auteur» ou du degré de fictionalité, la participation du spectateur s'avère cependant cruciale. La perspective que j'ai adoptée pour aborder l'univers diégétique, redevable des travaux menés depuis une vingtaine d'années sur les «mondes possibles» par les théoriciens de la littérature, se situe d'emblée du côté d'une pragmatique, puisqu'elle ne postule pas une «ontologie» des états

de choses proposés par le film, mais une «coopération textuelle» du spectateur (au sens d'Umberto Eco). Cette approche m'a permis notamment de saisir le rôle que peuvent jouer les sons acoustiques quant à l'accessibilité d'un monde à l'autre, et d'évaluer la place du «plan de référence» verbal dans la constitution d'un monde.

### LA POSTERITE DES ATTRACTIONS CHANTEES

En discutant les textes traitant de la notion d'«attraction» qui, selon moi, permet de rendre compte d'un aspect central de l'oralité, j'ai tenté de définir son emploi chez les théoriciens et les chroniqueurs d'époque, puis dans la perspective d'en faire une notion applicable au cinéma parlant. On a vu que si Gaudreault et Gunning n'ont jamais explicitement usé de cette notion pour aborder la voix, ils n'ont toutefois pas exclu cette possibilité. La voix-attraction se définit dans ce contexte historique comme un régime vocal qui, interagissant au sein d'un ensemble de pratiques spectaculaires, précède ou concurrence la standardisation des pratiques sonores. La dialectique entre oralité et fixation peut dès lors s'examiner à un niveau textuel, pour autant que l'on considère l'éventualité d'une voix pseudo-attractionnelle: l'attraction constitue le pôle inverse de l'«intégration». En raison de cette corrélation entre les caractéristiques du dispositif et celles du texte filmique, j'ai proposé d'utiliser le même terme pour convoquer deux niveaux: il y a d'une part l'intégration qui résulte de la fixation induite par le support utilisé, d'autre part celle qui s'applique aux différents paramètres liés à l'instauration d'une unité discursive. Cette dernière assure une lisibilité optimale et permet l'instauration d'une diégèse autonome de laquelle le monde du spectateur est implicitement forclos.

Comme l'a montré Édouard Arnoldy à propos des Phonoscènes, des «Filmparlants» et des bandes Vitaphone, les premières tentatives de films sonorisés dans les années 10 et 20 se sont accompagnées d'une hétérogénéité et d'une forme de spectacle qui présentent une parenté évidente avec le cinéma des premiers temps, celui-ci s'inscrivant, comme je pense l'avoir montré, dans des pratiques issues des spectacles vivants «de type vaudeville». Le début des années 30 est en outre marqué par une forte présence de la voix-attraction, caractéristique majeure de l'«interrègne» que

Noël Burch désignait comme une époque intermédiaire où l'apparition de la parole synchrone n'avait pas encore évincé toute trace d'un spectacle de type présentationnel. Pourtant, on peut dire que ce genre de spectacle se poursuit après les années 30 dans des films qui connaissent l'insertion de parties «chorégraphiées» au sens large (films d'action, d'arts martiaux, comédies musicales), mais aussi, ne serait-ce que ponctuellement, l'intervention de parties chantées, à l'instar des récents *8 Femmes* d'Ozon ou *Pas sur la bouche* de Resnais, qui renouent avec un sous-genre (le film-opérette) très prisé dans les premières années du parlant. En raison de ce type de films, j'ai proposé de corréler la valeur graduelle de l'attraction d'une séquence avec le degré d'intégration musicale du langage parlé: plus le chant s'impose avec force, plus les caractéristiques de la performance vocale étudiées par Zumthor sont exhibées, au détriment souvent de l'intelligibilité des paroles ou de leur fonction narrative.

De nos jours, le déclin de l'expression vocale «communautaire» au cinéma et de la comédie musicale (ou du western, qui donnait souvent lieu à des moments chantés en vertu de leur dimension folklorique et de leur rôle de cohésion sociale) pourrait laisser penser que les productions hollywoodiennes ont totalement évacué la voix-attraction. Néanmoins, on constate qu'elle se perpétue dans des films hybrides qui empruntent momentanément les atours de la comédie musicale. On peut penser en particulier aux films dont l'intrigue se déroule dans un milieu qui correspond à l'origine de l'usage attractionnel de la parole dont hérite le cinéma des premiers temps, comme *Moulin Rouge* (Baz Luhrman, 2001) ou *Chicago* (Bob Marshall, 2002). D'autres films comportent toutefois des parties chantées sans ressortir globalement à ce genre, à l'instar du dernier film de Tim Burton, *Charlie et la chocolaterie* (*Charlie and the Chocolate Factory*, 2005), *remake* du film homonyme des années 70 qui présentait déjà des moments dévolus au chant sur un mode proche du *Magicien d'Oz* (Victor Fleming, 1939, film où voix et Technicolor se combinent pour renforcer la valeur attractionnelle).

Dans le film de Burton, les paroles des chants entonnés par les «clones» lors de chorégraphies parodiques conservent par exemple un rapport étroit avec l'intrigue, les protagonistes principaux demeurant spectateurs de ces performances effectuées dans l'usine (sorte de «parc d'attractions»). En outre, la voix-attraction ne

domine que dans de brèves séquences, car *Charlie et la chocolaterie* se caractérise par un emploi diversifié de la voix qui témoigne de la légitimité des catégories que j'ai distinguées dans mon analyse des régimes vocaux. On y trouve en effet une voix synchrone d'automates parlants (le chant en chœur des poupées à l'entrée de l'usine), une voix déliée motivée par auricularisation interne (alors qu'un garçon prononce une phrase *in*, il est « doublé » par Christopher Lee qui interprète le père de Willy Wonka), une voix-over comme support d'un enchâssement (le récit du grand-père) et une autre qui apparaît dès l'ouverture du film pour lui donner une allure de conte, supposément extradiégétique, puis appréhendée rétrospectivement dans l'ultime plan du film comme provenant de l'un des « clones » de l'usine qui apparaît à l'avant-plan en tant que commentateur « juxtadiégétique ». *Charlie et la chocolaterie* prouve qu'en 2005, une grande production à succès recourant à des effets spéciaux numériques de pointe peut également exploiter le potentiel des différents régimes vocaux. Cette même année, Burton a d'ailleurs coréalisé *Les Noces funèbres* (*Corpse Bride*), un film d'animation qui présente également des moments chantés sur le modèle d'un autre long métrage scénarisé par le cinéaste, *L'Étrange Noël de M. Jack* (*Nightmare before Christmas*, Henry Selick, 1993). Dans ces deux films, le chant contribue de façon décisive à *donner vie* aux poupées filmées image par image, ce qui est en quelque sorte thématiquement sur le plan diégétique puisque plusieurs moments d'attraction sont inspirés de la danse macabre. On voit que la dominante attractionnelle de la voix n'a pas disparu avec la fin du *musical* hollywoodien.

#### LA VOIX-OVER AUJOURD'HUI

Bien qu'Albert Laffay déclare en 1964, presque vingt ans après avoir manifesté son enthousiasme pour les films américains comportant une voix-over, que « cette technique fut un peu une mode »<sup>824</sup>, la voix-over me paraît également avoir joué un rôle de premier plan à une époque plus récente de l'histoire du cinéma. Dans la France des années 60 émerge une « modernité cinématographique » (Varda, Resnais, Robbe-Grillet, Duras) et une « Nouvelle Vague » (Godard, Chabrol, Truffaut) dont la spécificité stylistique repose

824. A. Laffay, *Logique du cinéma*, op. cit., p. 81.

pour une part non négligeable sur certains usages inédits ou particuliers de la voix-over, pour les uns globalement définis par une «littérisation» du texte, pour les autres caractérisés par une grande liberté (voire une désinvolture) dans les effets de désancrage de la voix «dans» l'image. On rétorquera certes que ces films ont près de quarante ans. Ils n'en ont pas moins durablement marqué une certaine conception d'un cinéma d'«auteur» dont, justement, la voix auctoriale s'avère, depuis Welles, l'un des traits distinctifs. Notons par ailleurs qu'en octobre 2004 encore, une revue spécialisée comme *Vertigo* jugeait bon de publier un dossier consacré à la voix-over, dont le titre, «Voix off: Qui nous parle?», est révélateur du mystère dont n'ont cessé de se parer les voix sans corps<sup>825</sup>.

Quel que soit le genre filmique, il semble bien que la voix-over puisse y trouver sa place, tant dans les films rangés sous l'étiquette (certes fourre-tout, mais qui a une réalité dans le domaine de l'exploitation et de la diffusion) du «cinéma d'art et d'essai» (Alexandre Sokourov, Wong Kar-Wai ou Jean-Luc Godard, qui a fait de sa propre voix, dans ses films des années 90, un véritable signe de reconnaissance), dans les documentaires «auteorisés» de Moore à propos desquels on peut affirmer que le caractère charismatique de la voix-over du cinéaste est pour beaucoup dans leur succès (et dans leur démagogie) que dans les productions hollywoodiennes. Ainsi la quasi-totalité des longs métrages de fiction réalisés par Martin Scorsese<sup>826</sup> depuis 1990 comporte-t-elle une voix-over (et même trois en alternance dans *Casino*), alors que la voix du cinéaste apparaît dans ses documentaires «personnalisés» sur l'histoire du cinéma américain et italien. On pourrait voir dans ses films une résurgence non significative d'un certain classicisme, lié au genre du film noir ou du film de gangsters. Toutefois, la voix-over est également récurrente dans certaines réalisations «post-modernes» qui multiplient à outrance les citations et les jeux sur le récit. La propension de la voix-over à introduire des enchâssements y est exploitée à l'extrême. Pour ne citer que deux films qui connurent une diffusion très large, on peut mentionner *Le Fabuleux destin d'Amélie Poulain* (Jean-Pierre Jeunet, 2001) et la dernière réalisation en deux volets de Quentin Tarantino, *Kill Bill*

825. *Vertigo*, N° 26, octobre 2004.

826. Soit: *Les Affranchis* (*Goodfellas*, 1990); *L'Âge de l'innocence* (*The Age of Innocence*, 1993); *Casino* (1995); *À tombeau ouvert* (*Bringing out the Dead*, 1999); *Gangs of New York* (2002); *Les Infiltrés* (*The Departed*, 2006).

(2003 et 2004), où la voix-over permet de relier des séquences très hétérogènes (à l'exemple du passé d'une tueuse japonaise, raconté sous la forme d'un manga). Produit par le même Tarantino, la récente adaptation du *graphic novel* de Frank Miller, *Sin City* (Frank Miller et Roberto Rodriguez, 2005), techniquement fort complexe au niveau des retouches appliquées à l'image noir/blanc pour retrouver l'atmosphère urbaine surréaliste de l'univers graphique de Miller, est continuellement pourvue de voix-over, les personnages se relayant en une série de monologues intérieurs. Alors que les planches de Miller contiennent plutôt peu de verbal, il semble que les concepteurs du film aient voulu renvoyer globalement au médium BD en exhibant la présence du verbal (sous une forme orale qui actualiserait le « discours intérieur » du bédéphile).

Par ailleurs, l'un des principaux débouchés actuels des techniques informatiques de synthèse de l'image réside dans la réalisation de films d'animation coûteux et destinés à un public de tous âges. Or, dans ce genre de films on exploite à l'envi le potentiel humanisant de la voix, faisant des acteurs qui prêtent leur voix aux personnages un facteur déterminant de la promotion. Par-delà la facticité affichée du monde proposé à l'image, la voix, synchrone ou *over*, constitue un facteur décisif de proximité avec le public. Ce « réalisme » qui s'effectue *via* l'expression vocale est patent dans la séquence pré-générique des *Indestructibles* (*The Incredibles*, Brad Bird, 2004), dans laquelle les différents personnages de synthèse, nullement « photoréalistes », sont filmés en gros plan lorsqu'ils se confient en voix synchrone à la caméra dans un pastiche d'interview télévisée (l'image est d'ailleurs de format carré dans cette seule séquence afin de rappeler le format convoqué par cet intertexte), ou dans le film d'animation *Creature Comforts* de Nick Park (1989) composé de faux entretiens avec des animaux (dans lesquels l'effet comique repose sur leur humanisation outrancière), la perche et le micro étant présents « dans le champ » pour souligner le niveau (fictif) de la captation. Cet effet documentaire, totalement incongru dans des films possédant un si fort degré de fictivité – et de fictionalité pour le long métrage de Bird, puisque le film raconte l'histoire d'une famille de super-héros – met en évidence le rôle endossé par le grain de la voix dans les films d'animation.

Pour saisir plus précisément l'importance acquise par la voix-over dans la communication audiovisuelle, il faudrait sortir de l'« institution » du cinéma et s'interroger également sur d'autres

séries culturelles et technologiques que celles dont il a été question ici. On pourrait alors envisager, au-delà du principe de l'inscription phonographique, les développements plus récents qu'ont connus les techniques d'enregistrement et de diffusion de la voix, notamment avec l'apparition des supports magnétiques, puis numériques. Ce travail reste à faire, mon étude se contentant de fournir un cadre théorique qui devrait permettre d'appréhender certains de ces phénomènes qui, dans tous les cas, ressortissent à *l'écrit*.

### INTERNET: UN NOUVEL ESSOR POUR LA VOIX-OVER

Dans l'introduction de son ouvrage paru en 2004, Rodney Saulsberry, s'adressant aux acteurs qui se destinent à travailler «en voix-over», prétend qu'«aujourd'hui, il y a plus de débouchés professionnels pour la voix humaine qu'il n'y en a jamais eu»<sup>827</sup>. Cette affirmation n'est pas seulement rhétorique, même sous la plume d'un professionnel de la *captatio benevolentiae*. Si le statut social ou la légitimation culturelle d'un présentateur radio ne sont pas ceux d'un troubadour ou d'un griot, il faut bien reconnaître que l'essor formidable des technologies de la communication que nous avons connu depuis environ vingt ans n'est pas sans élargir la palette des manifestations *over* et des interactions entre l'expression humaine et les moyens mécaniques, électriques et informatiques. Le «retour en force de l'oralité primitive» dont Paul Zumthor rendait compte dans les années 80 en mentionnant la télévision et la radio est plus prégnant que jamais dans la vie quotidienne des «audiospectateurs» du début du XXI<sup>e</sup> siècle<sup>828</sup>.

L'existence aux États-Unis d'un véritable créneau éditorial constitué par les livres présentant des recettes aux futurs «acteurs de voix-over» est l'indice de l'omniprésence actuelle de cet usage de la voix. Les aspects abordés par les auteurs de ces guides, qui ont généralement fait eux-mêmes carrière dans le domaine, témoignent indiscutablement de l'importance de l'implantation de la voix-over, tant sur le plan institutionnel qu'au niveau des pratiques. Or il me semble que sur plusieurs points, les activités menées aujourd'hui par ces spécialistes reconduisent sur une base

827. Rodney Saulsberry, *You Can Bank on your Voice. Your Guide to a Successful Career in Voice-Overs*, Agoura Hills: Tomdor, 2004, p. 12.

828. P. Zumthor, «Le geste et la voix», *art. cit.*, p. 73.



technologique différente des principes et des domaines d'exercice qui constituent un prolongement des usages de la voix enregistrée de la fin du XIX<sup>e</sup> siècle, puis de l'époque des pionniers de la sonorisation au cinéma. Par exemple, la notice biographique d'Elaine A. Clark, fondatrice de l'école Voice One en 1986, mentionne sur la quatrième de couverture de son ouvrage que l'auteur « est une actrice de voix travaillant également devant la caméra, qui a été vue et entendue dans des centaines de publicités, de jeux vidéo, de sites Web, de jouets parlants et d'autres produits vocaux, dans des films industriels et dans des systèmes de messagerie vocale »<sup>829</sup>. Certains termes de cette énumération montrent que les « machines parlantes » qui ont abouti à la voix synchrone du cinéma parlant se sont multipliées grâce à l'intrusion de l'électronique dans le quotidien des habitants des pays industrialisés, qui, corrélativement, appelle une humanisation de la machine – en tant que lieu de la communication, la boîte e-mail simule l'oralité – ou son exhibition ludique (jouets et gadgets vocaux). En pourvoyant ces produits manufacturés, successeurs des automates parlants du XVIII<sup>e</sup> siècle, les fabricants confèrent, grâce à la reproduction de la voix, l'apparence du « vivant » à des appareils utilisés souvent en toute solitude. L'anthropomorphisation ne passe pas toujours par la synchronisation vocolabiale, les connotations humanisantes de la vocalité suffisant à faire l'intérêt du produit. On retrouve là le paradigme de l'« animation par la voix » duquel procèdent les utilisations de la voix-over lorsqu'il s'agit d'« incarner » un personnage en dépit de la nécessaire désincarnation du procédé. Dans l'ouvrage de James Alburger, véritable manuel d'entraînement des capacités vocales (il contient même un CD avec des exercices), un chapitre est consacré aux méthodes de personnalisation de la voix en vue d'« animer des personnages » (« *animation and character voices* »)<sup>830</sup>. Les conseils proposés se fondent notamment sur la traditionnelle analyse psychologique des personnages, et révèlent combien, ainsi que je l'ai indiqué dans mon schéma de l'énonciation, la vocalité peut renvoyer à la personnalisation d'un être fictionnel.

Sur des supports tels que le CD-ROM, les sites Web ou les jeux vidéo, la voix-over intervient soit sur le mode de l'adresse pour

829. Elaine A. Clark, *There's Money Where your Mouth is. An Insider's Guide to a Career in Voice-Overs*, New York: Back Stage Books, 2000.

830. James Alburger, *The Art of Voice Acting. The Craft and Business of Performing for Voice-Over*, Burlington: Focal Press, 2002, pp. 145-158.

exposer l'« attraction » interactive proposée par le produit, soit pour nous introduire dans le monde créé par (audio)visualisation. Ces mondes possibles ne sont pas nécessairement soumis à une organisation narrative – quoiqu'en anglais le mot « *narration* » se soit imposé quelle que soit l'organisation discursive du texte proféré, ce qui souligne combien cette notion tend à englober toute pratique « intégrative » – mais peuvent être constitués d'un ensemble de biens de consommation à travers lesquels la voix nous guide. Saulsberry décrit ainsi ce nouveau marché : « La voix des sites Web est rapidement devenue presque aussi populaire dans le domaine des affaires. Les grandes corporations cherchent des gens pour représenter leurs compagnies en guidant par la voix les utilisateurs sur leurs sites. »<sup>831</sup> La voix-over intervient alors pour augmenter la valeur informative du message en exploitant les processus de (co)référence, et pour dynamiser la lecture grâce aux variations que permet la vocalité. En se faisant l'émissaire des grandes firmes, la voix contribue à leur donner un visage plus humain parce qu'elle simule une communication interpersonnelle. Sur ces sites « parlés », la voix-over endosse la même fonction de construction d'un monde qu'au cinéma. On voit que la notion de « diégèse » peut s'appliquer au champ actuel de l'audiovisuel, et donc bien au-delà du canonique « film narratif de fiction ».

Ces ouvrages pédagogiques semblent montrer que les pratiques récentes suscitées par les nouvelles technologies demandent des compétences vocales qui, en elles-mêmes, diffèrent peu de ce que l'on exige de locuteurs *over* depuis l'essor de la radio à la fin des années 20. Ainsi les régimes vocaux identifiés dans ce travail permettent à mon sens de distinguer les principaux domaines d'application de la voix-over, ainsi que des procédés plus spécifiques à l'intérieur de ceux-ci : le doublage, la post-synchronisation de films d'animation et les boîtes vocales relèvent de la voix-action, l'animation radio et la publicité de la voix-attraction, les films industriels, CD-ROM et jeux vidéo de la voix-narration, alors que d'autres productions se situent entre ces deux dernières (les bandes-annonces, les sites Web, etc.). Les conseils pratiques témoignent du rôle central joué par la « performance vocale » au sens de Zumthor dans les milieux actuels de la publicité, de la communication, du *show-business* et de l'information, qui recourent

831. R. Saulsberry, *op. cit.*, p. 99.

massivement à l'audiovisuel. Les temps sont loin où le « message publicitaire » emblématique consistait en une affiche conjuguant du texte et une image fixe, comme cela était le cas à l'époque des textes sémiologiques de Roland Barthes. L'interaction entre les signes iconiques et verbaux n'en demeure pas moins le fondement de ce type de communication, même s'il paraît aujourd'hui indispensable d'inclure dans une conceptualisation de ces phénomènes la dimension orale issue de la tradition radiophonique et télévisuelle.

Dans mon ouvrage, je me suis principalement limité à un objet restreint, le médium « cinéma », considéré à quelques périodes-clés de l'évolution de sa composante sonore dans une volonté de discuter, dans le cadre d'une analyse de la place et du rôle de la voix humaine à l'intérieur de dispositifs audiovisuels engageant une « machinerie », certaines notions classiques en analyse du film (« narrateur », « énonciation » ou « diégèse »). La perspective intermédiaire et interdisciplinaire que j'ai tenté de développer entend contribuer à l'élaboration d'une approche théorique globale des phénomènes vocaux et verbaux dans leurs liens à l'image animée, et plus généralement à une réflexion sur cette intégration de « l'humain dans la machine » qu'induisent les moyens de communication de masse développés ou perfectionnés au cours du XX<sup>e</sup> siècle.



## BIBLIOGRAPHIE

ADAM Jean-Michel :

– *Le Récit*, Paris : PUF, 1984.

– « Si hypothétique et l'imparfait », in *Langue et littérature*, Paris : Hachette, 1992.

– *Les Textes : types et prototypes*, Paris : Nathan, 1992.

ADLER Laure, *Marguerite Duras*, Paris : Gallimard, 1998.

ADORNO Theodor W., Hanns EISLER, *Musique de cinéma*, Paris : L'Arche, 1972 [1947].

ALBERA François :

– « Introduction », in *Les Formalistes russes et le cinéma. Poétique du film*, Paris : Nathan, 1996.

– « Ouverture », 1895, N° 38 (« Musique! »), 2002.

– « Du bonimenteur à la voix off », *Vertigo*, N° 26, 2004.

ALBERA François, GAUDREULT André, « Apparition, disparition et escamotage du bonimenteur dans l'historiographie française du cinéma », in Giusy Pisano et Valérie Pozner (dir.), *Le Muet a la parole. Cinéma et performance à l'aube du XX<sup>e</sup> siècle*, Paris : AFRHC, 2005.

ALBERA François, TORTAJADA Maria, « L'Epistémè <1900> », in André GAUDREULT, Catherine RUSSELL et Pierre VÉRONNEAU (dir.), *Le Cinématographe, nouvelle technologie du XX<sup>e</sup> siècle*, Lausanne : Payot, 2004.

ALSBURGER James, *The Art of Voice Acting. The Craft and Business of Performing for Voice-over*, Burlington : Focal Press, 2002.

ALTHUSSER Louis, « Le « Piccolo », Bertolazzi et Brecht (Notes sur un théâtre matérialiste) », in *Pour Marx*, Paris : Maspero, 1975 [1962].

ALTMAN Rick :

– « Moving Lips: Cinema as Ventriloquism », *Yale French Studies*, N° 60, 1980.

– *La Comédie musicale hollywoodienne*, Paris : Armand Colin, 1992 [1987].

ALTMAN Rick (dir.), *Sound Theory, Sound Practice*, New York; London: Routledge, 1992.

– «Naissance de la réception classique. La campagne pour standardiser le son», *Cinémathèque*, N° 6, 1994.

– «Technologie et textualité de l'intermédialité», *Sociétés & représentations*, N° 9, 2000.

– *Silent Film Sound*, New York: Columbia University Press, 2004.

AMIEL Vincent, HERPE Noël, «C'est la parole qui incarne. Entretien avec Dominique Païni», *Double Jeu*, N° 3 («Sacha Guitry et les acteurs»), 2006.

ANDERSON Joseph L., RICHIE Donald, *The Japanese Film: Art and Industry*, Princeton: Princeton University Press, 1960.

ARLAUD Rodolphe-Maurice, *Cinéma-Bouffé. Le Cinéma et ses gens*, Paris: Éditions Jacques Melot, 1945.

ARNAUD Claude, «La voix de son maître», *Cinématographe*, N° 86 («Sacha Guitry»), février 1983.

ARNHEIM Rudolph, *Le Cinéma est un art*, Paris: L'Arche, 1989 [1933].

ARNOLDY Édouard:

– *Une histoire du cinéma parlant et sa diagonale mobile. De l'art cinématographique, des «films chantants et parlants» et de leurs séries*, thèse de doctorat, Université de Liège, 2001.

– «De la forme sérielle. L'entre du muet et du parlant: spectacle, chanson, cinéma et Cie», in Anna Antonini (dir.), *Il film e i suoi multipli*, Udine: Università degli studi di Udine, 2003.

– *Pour une histoire culturelle du cinéma. Au-devant de «scènes filmées», de «films chantants et parlants» et de comédies musicales*, Liège: Éditions du Céfal, 2004.

AVELINE Claude, *Chroniques d'un cinéophile*, Paris: Séguier, 1994.

BALAZS Béla, *Le Cinéma*, Paris: Payot 1979 [1949].

DE BANVILLE Théodore, *La Lanterne magique*, Paris: Charpentier, 1883.

BARJAVEL René, «Miracles à vendre», *Journal du Dimanche*, 4 septembre 1977.

BARNIER Martin, *En route vers le parlant. Histoire d'une évolution technologique, économique et esthétique du cinéma (1926-1934)*, Liège: Éditions du Céfal, 2002.

BARTHES Roland:

– «Rhétorique de l'image», *Communications*, N° 4, 1964.

– *Le Degré zéro de l'écriture*, Paris: Seuil, 1972 [1953].

- «Le grain de la voix», in *L'Obvie et l'obtus. Essais critiques III*, Paris: Seuil, 1982 [1972].
- *Poétique du récit*, Paris: Seuil, 1977.
- BAUDRY Jean-Louis, «Le dispositif: approches métapsychologiques de l'impression de réalité», *Communications*, N° 23, 1975.
- BELTON John, «Technology and Aesthetics of Film Sound», in Elisabeth Weis et John Belton (dir.), *Film Sound. Theory and Practice*, New York: Columbia University Press, 1985.
- BENAYOUN Robert, *Alain Resnais, arpenteur de l'imaginaire*, Paris: Stock Cinéma, 1980.
- BENJAMIN Walter, «Le conteur» [1936], in *Œuvres III*, Paris: Gallimard, 2000.
- BENVENISTE Émile, «Les relations de temps dans le verbe français» [1959], in *Problèmes de linguistique générale*, Vol. 1, Paris: Gallimard, 1966.
- BERG Chuck, ERSKINE Tom, *The Encyclopedia of Orson Welles*, New York: Facts on File, 2003.
- BERNARD André, GAUTEUR Claude (éd.), *Sacha Guitry. Le Cinéma et moi*, Paris: Ramsay, 1977.
- BLOM Ivo, VAN DOOREN Ine, «Ladies and Gentlemen, Hats off, Please!», *Iris*, N° 22, 1996.
- BOILLAT Alain:
- *Deux traductions de «La Chevelure de Bérénice» de Claude Simon. Paramètres pour l'analyse de traductions littéraires*, Lausanne: Centre de Traduction littéraire, 1999.
- *La Fiction au cinéma*, Paris: L'Harmattan, 2001.
- «Jean-Marie Schaeffer, Pourquoi la fiction? et Roger Odin, *De la fiction*», *Iris*, N° 30, 2004.
- «Le colloque *Le Muet a la parole* de l'auditorium du Louvre», 1895, N° 44, décembre 2004.
- *Les Voix du cinéma. Boniment, voix synchrone, déliaison et voix-over*, thèse de doctorat, Université de Lausanne, 2006.
- «*L'Ève future* et la série culturelle des «machines parlantes». Le statut singulier de la voix humaine au sein d'un dispositif audiovisuel», *Cinémas*, Vol. 17, N° 1, automne 2006.
- BOOTH William C., *The Rhetoric of Fiction*, Chicago: University of Chicago Press, 1961.
- BOTTOMORE Stephen, «Shots in the Dark. The Real Origin of Film Editing», in Thomas Elsaesser (dir.), *Early Cinema. Space, Frame, Narrative*, London: BFI, 1990.

BRANIGAN Edward, *Point of View in the Cinema*, Berlin; New York; Amsterdam: Mouton, 1984.

BRECHT Bertolt, *L'Achat du cuivre. Petit organon*, Paris: l'Arche, 1999 [1964].

BRINCKMANN Christine N., «Der Voice-Over im Film Noir» [1985], in *Die Anthropomorphe Kamera*, Zurich: Chronos, 1997.

BURCH Noël:

– *Pour un observateur lointain*, Paris: Gallimard; Éditions des Cahiers du cinéma, 1982 [1979].

– «Du muet, le parlant. Réflexions cursives sur un interrègne», in Christian BELAYGUE (dir.), *Le Passage du muet au parlant*, Toulouse: Cinémathèque de Toulouse, 1988.

– *La Lucarne de l'infini. Naissance du langage cinématographique*, Paris: Nathan, 1991.

– «Discipline et jouissance», in Bernard CHARDÈRE, Philippe DUJARDIN, Lucette LOMBARD-VALENTINO (dir.), *Lumière, le cinéma*, Lyon: Institut Lumière, 1992.

CAMPAN Véronique, *L'Écoute filmique. Écho du son en image*, Saint-Denis: PUV, 1999.

CANUDO Ricciotto, *L'Usine aux images*, Biarritz: Éditions Séguiet, 1995 [1927].

CARBINE Mary, «The Finest Outside the Loop»: Motion Picture Exhibition in Chicago's Black Metropolis, 1905-1928», *Camera obscura*, N° 23, 1990.

CARDINAL Serge, «Médiation ou modulation sonore?», *Cinémas*, Vol. 9, N° 1, 1999.

CARLIER Christophe, *Marguerite Duras. Alain Resnais. Hiroshima mon amour*, Paris: PUF, 1994.

CÉLINE Louis-Ferdinand, *Mort à crédit*, Paris: Denoël et Steele, 1936.

CENDRARS Blaise, *Dan Yack*, Paris: Denoël, 2002 [1929].

CHAPERON André, «L'institut de filmologie, une tentative d'interdisciplinarité», *Études de Lettres*, N° 2, 1993.

CHARDÈRE Bernard, Guy et Marjorie BORGÉ, *Les Lumière*, Lausanne: Payot, 1985.

CHARTIER Jean-Pierre, «Les films à la première personne et l'illusion de réalité au cinéma», *La Revue du Cinéma*, N° 4, 1947.

CHATEAU Dominique:

– «Diégèse et énonciation», *Communications*, N° 38, 1983.



– «Le rôle de la musique dans la définition du cinéma comme art: à propos de l'avant-garde des années 20», *Cinémas*, Vol. 3, N° 1, 1992.

CHÂTEAUVERT Jean:

– «Il faut trouver la voix», *Cinémas*, Vol. 3, N° 1, 1992.

– *Des mots à l'image. La Voix over au cinéma*, Québec; Paris: Nuit Blanche; Méridiens Klincksieck, 1996.

– «Le cinéma muait», *Iris*, N° 22, 1996.

– «Regards à la caméra: les comédiens en point de mire», in Michel MARIE et Laurent LE FORESTIER (dir.), *La Firme Pathé Frères 1896-1914*, Paris: AFRHC, 2004.

CHION Michel:

– *La Voix au cinéma*, Paris: Éditions des Cahiers du cinéma, 1982.

– *Le Son au cinéma*, Paris: Éditions des Cahiers du cinéma, 1985.

– *La Toile trouée*, Paris: Éditions des Cahiers du cinéma, 1988.

– *L'Audio-vision*, Paris: Nathan, 1990.

– *La Musique au cinéma*, Paris: Fayard, 1995.

– *Un art sonore, le cinéma*, Paris: Éditions des Cahiers du cinéma, 2003.

CLAIR René, *Comédies et commentaires*, Paris: Gallimard, 1959.

CLARK Elaine A., *There's Money Where your Mouth is. An Insider's Guide to a Career in Voice-Overs*, New York: Back Stage Books, 2000.

CŒUROY André, CLARENCE G., *Le Phonographe*, Paris: Éditions Kra, 1929.

COHEN Keith, «Pleasures of Voicing: Oral Intermittences in Two Films by Alain Resnais», *L'Esprit créateur*, Vol. 30, N° 2, 1990.

COISSAC G.-M., *Manuel pratique du conférencier-projectionniste*, Paris: La Bonne Presse, 1908.

COLIN Michel, *Langue, film, discours*, Paris: Klincksieck, 1985.

COLPI Henri, «Musique d'Hiroshima», *Cahiers du Cinéma*, N° 103, 1960.

COMOLLI Jean-Louis:

– «Caméra, perspective, profondeur de champ, II (5)», *Cahiers du Cinéma*, N° 234-235, 1971.

– «Film/politique (2). *L'Aveu*: 15 propositions», *Cahiers du Cinéma*, N° 224, 1970.

COSANDEY Roland, «L'abbé Joye, une collection, une pratique», in Roland COSANDEY, André GAUDREAU et Tom GUNNING (dir.), *Une invention du diable? Cinéma des premiers temps et religion*, Sainte-Foy; Lausanne: Université Laval; Payot, 1992.

CRANGLE Richard, «Next Slide Please: The Lantern Lecture in Britain,

1890-1910», in Richard ABEL et Rick ALTMAN (dir.), *The Sounds of Early Cinema*, Bloomington; Indianapolis: Indiana University Press, 2001.

DÄLLENBACH Lucien, *Le Récit spéculaire*, Paris: Seuil, 1977.

DANEY Serge, «L'orgue et l'aspirateur (Bresson, le Diable, la voix *off* et quelques autres)», in *La Rampe. Cahier critique 1970-1982*, Paris: Éditions des Cahiers du cinéma; Gallimard, 1983.

DE FRANCE Claudine, «Image et commentaire: du montré à l'évoqué», *Hors Cadre*, N° 3, 1985.

DELACOMMUNE Charles, «Le Synchronisme de la parole, de la musique et des bruits au cinéma», in Maurice NOVERRE, *Septumia. Légende-mime (Mœurs antiques)*, Bordeaux; Paris: Bourlange et Mollat; Dion, 1923.

DELEUZE Gilles, *Cinéma II. L'Image-temps*, Paris: Minuit, 1985.

DESANTI Dominique, *Sacha Guitry. 50 ans de spectacle*, Paris: Grasset; Fasquelle, 1982.

DESLANDES Jacques, *Histoire comparée du cinéma*, tome I, Paris: Casterman, 1966.

DEVAUX Frédérique, *Le Cinéma Lettriste (1951-1991)*, Paris: Éditions Paris Expérimental, 1992.

DOANE Mary Ann, «The Voice in the Cinema: The Articulation of Body and Space», *Yale French Studies*, N° 60, 1980.

DOYLE Arthur Conan, *Histoire du spiritisme*, Paris: Éditions du Rocher, 1981 [1927].

DUCHARTRE Pierre Louis, SAULNIER René, *L'Imagerie populaire. Les Images de toutes les provinces françaises du XV<sup>e</sup> siècle au Second Empire. Les plaintes, contes, chansons, légendes qui ont inspiré les imagiers*, Paris: Librairie de France, 1925.

DURRER Sylvie, *Le Dialogue romanesque*, Genève: Librairie Droz, 1994.

ECO Umberto:

– «Sémiologie des messages visuels», *Communications*, N° 15, 1970.

– *Lector in fabula. Le Rôle du lecteur*, Paris: Grasset, 1985 [1979].

– *Les Limites de l'interprétation*, Paris: Grasset, 1992.

EIKHENBAUM Boris:

– «Problèmes de ciné-stylistique» [1927], in François ALBERA (dir.), *Les Formalistes russes et le cinéma. Poétique du film*, Paris: Nathan, 1996.

– «Le mot et le cinéma» [1927], in François ALBERA (dir.), *Les Formalistes russes et le cinéma. Poétique du film*, Paris: Nathan, 1996.

EISENSCHITZ Bernard, «Ceux de chez nous», in *Sacha Guitry, cinéaste*, Locarno; Crisnée: Éditions du Festival de Locarno; Yellow Now, 1993.

EZRA Elisabeth, *Georges Méliès. The Birth of the Auteur*, Manchester; New York: Manchester University Press, 2000.

FIRK M., «*Hiroshima mon amour*, film scandaleux?», *Les Lettres françaises*, N° 773, 1959.

FÓNAGY Ivan, *La Vive voix*, Paris: Seuil, 1983.

FOREST Claude, *Les Dernières Séances. Cent ans d'exploitation des salles de cinéma*, Paris: CNRS, 1995.

FRANK Nino, «Un nouveau genre «policier». L'aventure criminelle», *L'Écran français*, N° 61, 28 août 1946.

FRESNAULT-DERUELLE Pierre, «Du linéaire au tabulaire», *Communications*, N° 24, 1976.

FREUD Sigmund, «L'inquiétante étrangeté» [1919], in *L'Inquiétante étrangeté et autres essais*, Paris: Gallimard, 1989.

GARDIES André:

– *Le Récit filmique*, Paris: Hachette, 1993.

– «Quand la narrativité vient aux films: les vues Lumière», in Jean ARROUYE, Marie-Claude TARANGER (dir.), *Rencontres. Croisements. Emprunts. Méthodologies de l'analyse d'images*, Aix-en-Provence: Publications de l'Université de Provence, 1996.

GARRONI Emilio:

– *Progetto di semiotica*, Bari: Éditions Laterza, 1972.

– «Langage verbal et éléments non-verbaux dans le message filmico-télévisuel», in Dominique NOGUEZ (dir.), *Cinéma: théorie, lectures*, Paris: Klincksieck, 1973.

GAUDREULT André:

– «Les aventures d'un concept: la narrativité», in Michel MARIE et Marc VERNET (dir.), *Christian Metz et la théorie du cinéma*, Paris: Méridiens Klincksieck, 1990.

– «Fonctions et origines du bonimenteur du cinéma des premiers temps», *Cinémas*, Vol. 4, N° 1, 1993.

– «Le retour du [bonimenteur] refoulé... (ou serait-ce le bonisseur-conférencier, le commentateur, le conférencier, le présentateur ou le «speaker»?)», *Iris*, N° 22, 1996.

– «Les vues cinématographiques selon Georges Méliès, ou: comment Mitry et Sadoul avaient peut-être raison d'avoir tort (même si c'est surtout Deslandes qu'il faut lire et relire)...», in Jacques MALTHÊTE et Michel MARIE

(dir.), *Georges Méliès, l'illusionniste fin de siècle?* Paris: Sorbonne Nouvelle, 1997.

– *Du littéraire au filmique*, Québec; Paris: Nota Bene, Armand Colin, 1999 [1988].

GAUDREULT André, LACASSE Germain, «L'écran ventriloque», in André GAUDREULT (dir.), *Au pays des ennemis du cinéma... Pour une nouvelle histoire des débuts du cinéma au Québec*, Montréal: Nuit Blanche, 1996.

GAUDREULT André, MARION Philippe, «Du sable dans les rouages du dispositif», in André GAUDREULT, Catherine RUSSELL et Pierre VÉRONNEAU (dir.), *Le Cinématographe, nouvelle technologie du XX<sup>e</sup> siècle*, Lausanne: Payot, 2004.

GEDULD Harry M., *The Birth of the Talkies. From Edison to Jolson*, Bloomington; London: Indiana University Press, 1975.

GENETTE Gérard:

– *Figures II*, Paris: Seuil, 1969.

– *Figures III*, Paris: Seuil, 1972.

– *Nouveau discours du récit*, Paris: Seuil, 1983.

GOFFMAN Erwin, *Façons de parler*, Paris: Minuit, 1987 [1981].

GORKI Maxime, «Le cinématographe Lumière», 6 juillet 1896 [trad. française de Valérie Pozner], 1895, N° 50, décembre 2006.

GOUDET Stéphane (dir.), *Positif, revue de cinéma. Alain Resnais*, Paris: Gallimard, 2002.

GREENBERG Larry, *The Benshi – Japanese Silent Film Narrators*, Tokyo: Urban Connections, 2001.

GUIDO Laurent, *L'Âge du rythme*, thèse de doctorat, Université de Lausanne, 2004.

GUITRY Sacha, *Cinquante ans d'occupations*, Paris: Presses de la Cité, 1993 [1947].

GUNNING Tom:

– «The Scene of Speaking: Two Decades of Discovering the Film Lecturer», *Iris*, N° 27, 1999.

– «Fantasmagorie et fabrication de l'illusion: pour une culture optique du dispositif cinématographique», *Cinémas*, Vol. 14, N° 1, 2003.

– «Pathé and the Cinematic Tale: Storytelling in Early Cinema», in Michel MARIE et Laurent LE FORESTIER (dir.), *La Firme Pathé Frères 1896-1914*, Paris: AFRHC, 2004.

HAMBURGER Käte, *Logique des genres littéraires*, Paris: Seuil, 1986 [1977].

HENDERSON Brian, «Tense, Mood, and Voice in Film (Notes after Genette)», *Film Quarterly*, Vol. 36, N° 4, 1983.

HEDIGER Vinzenz, *Verführung zum Film*, Marburg: Schüren Verlag, 2001.  
– «Putting the Spectators in a Receptive Mood», in Veronica INNOCENTI et Valeria RE (dir.), *Limina/Le Soglie del film*, Udine: Università degli studi di Udine, 2004.

HENDRYKOWSKA Malgorzata, «Professeur, acteur ou bouffon? Le bonimenteur sur les territoires polonais avant 1914», *Iris*, N° 22, 1996.

HOECK Leo H., *La Marque du titre. Dispositifs sémiotiques d'une pratique textuelle*, La Haye; Paris: Mouton, 1981.

HOFFMANN Ernst Theodor Amadeus, «Die Automate», in *Die Serapiens-Brüder*, tome 1, Darmstadt: Wissenschaftliche Buchgesellschaft, 1968 [1813].

HUSTER Francis, «Le Roman d'un tricheur», *L'Avant-scène théâtre*, N° 1181, 2005.

ICART Roger, *La Révolution du parlant, vue par la presse française*, Perpignan: Institut Jean Vigo, 1988.

ILLICH Ivan, *Du lisible au visible: la naissance du texte*, Paris: Éditions du Cerf, 1991.

JAKOBSON Roman:

– «Aspects linguistiques de la traduction» [1959], in *Essais de linguistique générale*, Paris: Minuit, 1963.

– «Linguistique et poétique» [1960], in *Essais de linguistique générale*, Paris: Minuit, 1963.

JOLY Martine, *L'Image et les signes*, Paris: Nathan, 1994.

JOST François:

– *L'Œil-caméra*, Lyon: PUL, 1987.

– *Un monde à notre image*, Paris: Méridiens Klincksieck, 1992.

– «Les mots pour le voir», in Francesco PITASSIO et Leonardo QUARESIMA (dir.), *Scrittura e immagine*, Udine: Università degli Studi di Udine, 1998.

– «The Voices of Silence», in Richard ABEL et Rick ALTMAN (dir.), *The Sounds of Early Cinema*, Bloomington; Indianapolis: Indiana University Press, 2001.

JULLIER Laurent:

– *Les Sons au cinéma et à la télévision*, Paris: Armand Colin, 1995.

– *L'Écran post-moderne*, Paris: L'Harmattan, 1997.

– *Cinéma et cognition*, Paris: L'Harmattan, 2002.

– *Star Wars. Anatomie d'une saga*, Paris: Armand Colin, 2005.

KEIM Jean A., *Un nouvel art. Le Cinéma sonore*, Paris: Albin Michel, 1947.

KEIT Alain, *Le Cinéma de Sacha Guitry. Vérités, représentations, simulacres*, Liège: Céfal, 1999.

KITTLER Friedrich, *Film, Grammophon, Typewriter*, Berlin: Brinkmann & Bose, 1986.

KOLB Bryan, WISHAW Ian Q., *Cerveau et Comportement*, Bruxelles: De Boeck Université, 2002.

KOMATSU Hiroshi, LONDEN Frances, «Mastering the Mute Image: The Role of the *Benshi* in Japanese Cinema», *Iris*, N° 22, 1996.

KOZLOFF Sarah, *Invisible Storytellers. Voice-over Narration in American Fiction Film*, Berkeley; Los Angeles; London: University of California Press, 1988.

KRACAUER Siegfried, *Theory of Film, The Redemption of Physical Reality*, London: Oxford University Press 1960.

KURODA Sige-Yuki, «Réflexions sur les fondements de la théorie de la narration», in Julia KRISTEVA, Jean-Claude MILNER, Nicolas RUWET (dir.), *Langue, discours, société*, Paris: Seuil, 1975.

LACASSE Germain:

– «De Passions en passions: le cinéma des débuts au Québec», in Roland COSANDEY, André GAUDREULT, Tom GUNNING (dir.), *Une invention du diable? Cinéma des premiers temps et religion*, Sainte-Foy; Lausanne: Université Laval; Payot, 1992.

– «Palimpsestes, trame et traces: tradition orale et notion de genre dans le cinéma des premiers temps», *Iris*, N° 20, 1995.

– «Du cinéma oral au spectateur muet», *Cinémas*, Vol. 9, N° 1, 1998.

– *Le Bonimenteur de vues animées. Le cinéma «muet» entre tradition et modernité*, Québec; Paris: Nota Bene, Méridiens Klincksieck, 2000.

LAFFAY Albert:

– «Le récit, le monde et le cinéma», *Les Temps modernes*, N° 21, mai-juin 1947.

– *Logique du cinéma*, Paris: Masson et C<sup>ie</sup>, 1964.

LARDEAU Yann, «Les pages arrachées au Livre de Satan», in *Sacha Guitry, cinéaste*, Locarno; Crisnée: Éditions du Festival de Locarno; Yellow Now, 1993.

LASTRA James, *Sound Technology and The American Cinema. Perception, Representation, Modernity*, New York; Chichester: Columbia University Press, 2000.

- LE GUAY Philippe, «Le double boîteux», *Cinématographe*, N° 86, 1983.
- LE MEN Ségolène, *Lanterne magique, tableau transparent*, Paris: Réunion des Musées nationaux, 1995.
- LEUTRAT Jean-Louis:  
 – *Hiroshima mon amour*, Paris: Nathan, 1994.  
 – *Vie des fantômes. Le fantastique au cinéma*, Paris: Éditions des Cahiers du cinéma, 1995.
- L'HERBIER Marcel, *Intelligence du cinématographe*, Paris: Éditions Buchet-Chastel, 1977 [1946].
- LIANDRAT-GUIGUES Suzanne, LEUTRAT Jean-Louis, *Alain Resnais. Liaisons secrètes, accords vagabonds*, Paris: Éditions des Cahiers du cinéma, 2006.
- LION Roger, «En avant la musique!», *Cinémagazine*, N° 47, 19 novembre 1926.
- LONDON Kurt, *Film Music*, London: Faber & Faber, 1936.
- LOTMAN Iouri, *Sémiotique et esthétique du cinéma*, Paris: Éditions sociales, 1977.
- MADIS Alex, *Sacha*, Paris: Éditions de l'Élan, 1950.
- MALTHÊTE Jacques:  
 – (dir.), *Essai de reconstitution du catalogue de la Star-Film*, Paris: CNC, 1981.  
 – «Méliès et le conférencier», *Iris*, N° 22 1996.
- MALTHÊTE-MÉLIÈS Madeleine, *Méliès l'enchanteur*, Paris, Hachette, 1973.
- MANDELSTAMM Valentin, «L'Avènement des Films parlants et synchronisés», *Cinémagazine*, N°s 5, 6, 7 et 8, 1<sup>er</sup> février, 8 février, 15 février et 22 février 1929.
- MANNONI Laurent, *Le Grand Art de la lumière et de l'ombre*, Paris: Nathan, 1995.
- MARIE Michel, «Un film sonore, musical, parlant», in Claude BAIBLÉ, Michel MARIE et Marie-Claire ROPARS-WUILLEUMIER (dir.), *Muriel. Histoire d'une recherche*, Paris: Éditions Galilée, 1974.
- MARTINET André, *Éléments de linguistique générale*, Paris: Armand Colin, 1970.
- MASSON Alain, *L'Image et la parole. L'Avènement du cinéma parlant*, Paris: La Différence, 1989.
- MATLOCK Jann, «Voir aux limites du corps: fantasmagories et femmes invisibles dans les spectacles de Robertson», in LE MEN (dir.), *Lanterne magique, tableau transparent*, Paris: Réunion des Musées nationaux, 1995.

MENY Jacques, «Méliès imaginé ou images de Méliès au cinéma et à la télévision», in Jacques MALTHÈTE et Michel MARIE (dir.), *Georges Méliès, l'illusionniste fin de siècle?* Paris: Sorbonne Nouvelle, 1997.

MERCIER Denis, (dir.), *Le Livre des techniques du son*, tome 3 («L'exploitation»), Paris: Éditions Fréquences, 1993 [1987].

MERIGEAU Pascal, *Le Nouvel Observateur*, N° 2113, 5 mai 2005.

METZ Christian:

- *Essais sur la signification au cinéma*, tome 1, Paris: Klincksieck, 1966.
- *Langage et cinéma*, Paris: Larousse, 1971.
- *Essais sur la signification au cinéma*, tome 2, Paris: Klincksieck, 1972.
- *Essais sémiotiques*, Paris: Klincksieck, 1977 [1974].
- *L'Énonciation impersonnelle ou le site du film*, Paris: Méridiens Klincksieck, 1991.
- *Le Signifiant imaginaire*, Paris: Christian Bourgois, 1993 [1977].

MEUSY Jean-Jacques, *Paris-Palaces, le temps des cinémas (1894-1918)*, Paris: CNRS, 1995.

MILNER Jean-Claude, *Ordres et raisons de langue*, Paris: Seuil, 1982.

MITRY Jean:

- *Esthétique et psychologie du cinéma*, Vol. 2, Paris: Éditions Universitaires, 1965.
- *La Sémiologie en question*, Paris: Cerf, 1987.

MONCLAIR Didier, «Une invention française. Le Ciné-pupitre Delacomune», *Cinémagazine*, N° 11, 16 mars 1923.

MONTALBETTI Christine, *La Fiction*, Paris: Flammarion, 2001.

MORIN Edgar, *Le Cinéma ou l'homme imaginaire*, Genève: Gonthier, 1965.

MOUREN Yannick, *Le Flash-Back*, Paris: Armand Colin, 2005.

MULLER Marcel, *Les Voix narratives dans «La Recherche du temps perdu»*, Genève: Droz, 1965.

MURCIA Claude, *Nouveau roman, nouveau cinéma*, Paris: Nathan, 1998.

MUSSER Charles:

- *The Emergence of Cinema: The American Screen to 1907*, New York; Toronto: Scribner's Sons; Collier Marmillon, 1990.
- «The Travel Genre in 1903-1904. Moving Towards Fictional Narratives», in Thomas ELSAESSER (dir.), *Early Cinema. Space, Frame, Narrative*, London: BFI, 1990.
- «Pour une nouvelle approche du cinéma des premiers temps: le cinéma d'attractions et la narrativité», in Jean GILI, Michèle LAGNY, Michel MARIE et Vincent PINEL (dir.), *Les Vingt Premières Années du cinéma français*, Paris: AFRHC, 1995.



NEDELEC Anne-Sophie, «L'auteur en voix *off*», in Francis VANOYE (dir.), *Cinéma et littérature*, Paris: Université de Paris, 1999.

NINEY François, *L'Épreuve du réel à l'écran*, Bruxelles: De Boeck, 2000.

NOGUEZ Dominique, *Éloge du cinéma expérimental*, Paris: Centre Georges Pompidou, 1979.

ODIN Roger:

– «L'entrée du spectateur dans la fiction», in Jacques AUMONT et Jean-Louis LEUTRAT J.-L. (dir.), *Théorie du film*, Paris: Albatros, 1980.

– *Cinéma et production de sens*, Paris: Armand Colin, 1990.

– (dir.), *L'Âge d'or du documentaire. Europe: Années cinquante*, tome 1, Paris: L'Harmattan, 1998.

ONG Walter, *The Presence of the Word*, New Haven; London: Yale University Press, 1967.

PANOFSKY Erwin, «Style et matière du septième art» [1936], in *Trois essais sur le style*, Paris: Gallimard, 1996.

PARCHOMENKO Laure, «Le fonds d'archives de la société Cinéas», 1895, N° 43, 2004.

PASOLONI Pier Paolo, «La langue écrite de la réalité» [1966], in *L'Expérience hérétique*, Paris: Payot, 1976.

PASSEK Jean-Loup, *Dictionnaire du cinéma*, Paris: Larousse, 1995.

PAVEL Thomas, *Univers de la fiction*, Paris: Seuil, 1986.

PERCHERON Daniel, «Le son au cinéma dans ses rapports à l'image et à la diégèse», *Ça Cinéma*, 1<sup>re</sup> année, 2<sup>e</sup> état, octobre 1973.

PERRIAULT Jacques, *Mémoires de l'ombre et du son. Une archéologie de l'audio-visuel*, Paris: Flammarion, 1981.

PEYTARD Jean, *Voix et traces narratives chez Stendhal*, Paris: Les Éditeurs Français Réunis, 1980.

PINGAUD Bernard:

– «Alain Resnais», *Premier Plan*, N° 18, 1961.

– «Nouveau roman et nouveau cinéma», *Cahiers du Cinéma*, N° 185, 1966.

PIRANDELLO Luigi, «Se il film parlante abolirà il teatro» [1929], cité par Alberto BOSCHI, «Il cinema sonoro nella teoria classica», in *L'Immagine acustica II, Cinegrafie*, N° 6, 1993.

PISANO Giusy:

– *Interférences entre son et image: un siècle de recherches avant le cinéma parlant*, thèse de doctorat, Paris, 1999.

- *Une archéologie du cinéma sonore*, Paris: CNRS, 2004.
- «Les spectacles mixtes: tradition ou anachronisme? Survivances sonores et visuelles de Robertson à Georges Lordier», in Giusy PISANO et Valérie POZNER (dir.), *Le Muet a la parole. Cinéma et performance à l'aube du XX<sup>e</sup> siècle*, Paris: AFRHC, 2005.

PORTES Jacques, *De la scène à l'écran. Naissance de la culture de masse aux États-Unis*, Paris: Belin, 1997.

POTTIER Bernard, (dir.), *Le Langage*, Paris: Centre d'étude et de promotion de la lecture, 1973.

POUILLON Jean, *Temps et roman*, Paris: Gallimard, 1946.

POZNER Valérie:

- «La ciné-déclamation en Russie», in Giusy PISANO et Valérie POZNER (dir.), *Le Muet a la parole. Cinéma et performance à l'aube du XX<sup>e</sup> siècle*, Paris: AFRHC, 2005.
- «Le bonimenteur <rouge>. Retour sur la question de l'oralité à propos du cas soviétique», *Cinémas*, Vol. 14, N° 2-3, 2005.

PREDAL René, *Études cinématographiques*, N° 64-68 («Alain Resnais»), 1968.

PRÉVOST Jean, «L'étudiant de Prague; Ombres blanches», *La Nouvelle Revue Française*, tome XXXII, 1929.

PRIEUR Jérôme:

- *Séance de lanterne magique*, Paris: Gallimard, 1985.
- *Le Spectateur nocturne*, Paris: Éditions de l'Étoile; Cahiers du cinéma, 1993.

PROUST Marcel, *Du côté de chez Swann*, Paris: Le Livre de Poche, 1992 [1913].

RAYNAULD Isabelle, «Les scénarios de la Passion selon Pathé (1902-1914)», *Une invention du diable? Cinéma des premiers temps et religion*, Sainte-Foy; Lausanne: Université Laval; Payot, 1992.

RICCEUR Paul:

- *Temps et récit*, tome 2, Paris, Seuil, 1984.
- *Temps et récit*, tome 3, Paris, Seuil, 1985.

ROBBE-GRILLET Alain, *L'Année dernière à Marienbad*, Paris: Éditions «J'ai lu», 1988 [1961].

ROPARS Marie-Claire:

- *De la littérature au cinéma. Genèse d'une écriture*, Paris: Armand Colin, 1970.
- *Le Texte divisé*, Paris: PUF, 1981.
- *Écraniques*, Lille: Presses Universitaires de Lille, 1990.

RYAN Marie-Laure, «Fiction, non-factuals, and the Principle of Minimal Departure», *Poetics*, N° 9, 1980.

SADOUL George, *Histoire générale du cinéma*, tome 1, Paris: Denoël, 1948.

SALT Barry, *Film Style and Technology: History and Analysis*, London: Starword, 1983.

SANAKER John Kristian, «La prose au cinéma ou Quand la voix *off* impose son rythme à l'image», in F. VANOYE (dir.), *Cinéma et littérature*, Paris: Université de Paris X, 1999.

SAULSBERRY Rodney, *You Can Bank on your Voice. Your Guide to a Successful Career in Voice-Overs*, Agoura Hills: Tomdor, 2004.

SCHAEFFER Jean-Marie:

– *L'Image précaire*, Paris: Seuil, 1987.

– *Pourquoi la fiction?* Paris: Seuil, 1999.

SICLIER Jacques, «Conversation avec Alain Resnais. «Je ne fais pas un film consciemment», *Le Monde*, 22 mai 1980.

SILVERMAN Kaja, *The Acoustic Mirror, The Female Voice in Psychoanalysis and Cinema*, Bloomington: Indianapolis, Indiana University Press, 1988.

SIROIS-TRAHAN Jean-Pierre, «La bataille du parlant, rue Sainte-Catherine, vingt ans avant le *Jazz Singer*», in André GAUDREAU (dir.), *Au pays des ennemis du cinéma... Pour une nouvelle histoire des débuts du cinéma au Québec*, Montréal: Nuit Blanche, 1996.

SOPOCY Martin, «Un cinéma avec narrateur. Les premiers films narratifs de James A. Williamson», *Cahiers de la Cinémathèque*, N° 29, 1979.

SOURIAU Anne (dir.), *Vocabulaire d'esthétique*, Paris: PUF, 1990.

SOURIAU Étienne, «L'impression de réalité au cinéma: les phénomènes de croyance», in Étienne SOURIAU (dir.), *L'Univers filmique*, Paris: Flammarion, 1953.

SUHAMY Henri, *Les Figures de style*, Paris: PUF, 1981.

THOMAS François, *L'Atelier Alain Resnais*, Paris: Flammarion, 1990.

TODOROV Tzvetan, *Littérature et signification*, Paris: Larousse, 1967.

TOULET Emmanuelle, *Cinématographe, invention du siècle*, Paris: Gallimard; Réunion des Musées nationaux, 1988.

TOULET Emmanuelle, BELAYGUE Christian (dir.), *Musique d'écran. L'Accompagnement musical du cinéma muet en France 1918-1995*, Paris: Réunion des Musées nationaux, 1994.

TRUFFAUT François :

- «Du cinéma pur», *Cahiers du Cinéma*, N° 70, avril 1957.
- «Préface», in *Sacha Guitry. Le Cinéma et moi*, Paris: Ramsay, 1977.

VAN BEUSEKOM Ansje, «The Rise and Fall of the Lecturer as Entertainer in the Netherlands», *Iris*, N° 22, 1996.

VAN ROSSUM-GUYON Françoise, «Point de vue ou perspective narrative», *Poétique*, N° 4, 1970.

VANOYE Francis, *Récit écrit, récit filmique*, Paris: Nathan, 1989.

VERNET Marc, *Figures de l'absence. De l'invisible au cinéma*, Paris: Éditions des Cahiers du cinéma, 1988.

VOGELEER Svetlana, «La relation point de vue et son application aux phrases existentielles initiales», in Walter DE MULDER, Franc SCHUEREWEGEN et Liliane TASMOWSKI (dir.), *Énonciation et parti pris*, Amsterdam; Atlanta: Rodopi, 1992.

VYGOTSKI Lev Semenovic, *Pensée et langage*, Paris: Éditions sociales, 1985.

WEINRICH Harald, *Le Temps*, Paris: Seuil, 1964.

WHISSEL Kristen, «Placing the Spectator on the Scene of History: the Battle Re-Enactment at the Turn of the Century, from Buffalo Bill's Wild West to the Early Cinema», *Historical Journal of Film, Radio and Television*, Vol. 22, N° 3, 2002.

WILLIAMS Alan, «Is Sound Recording Like a Language?», *Yale French Studies*, N° 60, 1980.

WILSON Emma, *Alain Resnais*, Manchester; New York: Manchester University Press, 2006.

WOLFE Charles :

- «On the Track of the Vitaphone Short», in M. BANDY (dir.), *The Dawn of Sound*, New York: MoMA, 1989, repris in Rick ALTMAN (dir.), *L'Immaginer acustica II. Il Passaggio dal muto al sonoro in America (Cinegrafie, N° 6)*, Ancona: Transeuropea, 1993.
- «Vitaphone Shorts and *The Jazz Singer*», *Wide Angle*, Vol. 12, N° 3, juin 1990.

WYSS André, *Éloge du phrasé*, Paris: PUF, 1999.

ZUMTHOR Paul :

- *Introduction à la poésie orale*, Paris: Seuil, 1985.
- «Le geste et la voix», *Hors Cadre*, N° 3 («Voix off»), 1985.
- *La Lettre et la voix*, Paris: Seuil, 1987.
- *Performance, réception, lecture*, Longueuil: Les Éditions du Préambule, 1990.

# INDEX NOTIONNEL

Acousmatique, acousmatisation: 25, 49, 52, 327, 342, 373, 385, 417, (note 736)-418, 473, 489-491.

Acousmètre: 417, note 476.

Adresse: 46-47, 51, 78-79, 106-107, 111, 120-121, 200, 205, 210-211, 220, 226, 232, 239-240, 252-259, 283-284, 293, 297-307, 427, 462, 497-498.

Arché (savoir de l'): 384-385, 426-428, 488.

Attestation (d'homodiégéticité): 437, 465.

Attraction-s: 26-27, 47, 51, 62, 69, 77-78, 102, 129-131, 152-153, 157, 167, 190, 197-308, 334, 420, 426-427, 485, 491-493.

Auteur: 113, 185, 282-284, 316.

Bande-annonce: 295-296, 310-313, 498.

*Benshi*: 34, 50, 59, 101, 107, 119-120, 130, 134, 176, 212.

Boniment-eur: 32-163, 166-167, 169, 171, 173, 177, 180-182, 184, 192, 194, 199, 201-202, 209, 211-212, 215, 224-225, 227, 234, 242-249, 251-253, 256, 258, 263, 273-274, 283-285, 288, 305, 326, 340, 345, 390, 400, 425-426, 438, 445, 483, 487-488.

Bonisseur, aboyeur: note 30, 33, 44, 78, 101, 121, 150, 153, 241, 244, 256, 306.

Bruit du projecteur: 103, 130-148, 150-151, 413.

Bruits: 17, 28-30, 79, 87, 136, 172, 204, 248, 302, 406, 411, 420, 428.

Chant: 59, 77-78, 86-87, 132-133, 138, 141, 146, 211, 218, 220-221, 227-228, 232-238, 270, 275, 277-278, 280-281, 284-287, 292-293, 296-297, 311, 334, 361, 393, 449, 492-493, note 86.

Comédie musicale (*musical*): 220, 227-235, 238, 426-427, 492.

Communion, communautaire: 102-103, 109, 159, 182, 232-233, 284-285.

Conférence, conférencier (lorsque le terme est distingué de boniment/bonimenteur): 33, 41, 44-46, 50-51, 62-82, 88, 95, 99-101, 106-107, 109, 119-120, 122, 128-130, 134, 136, 173-180, 213-215, 243-245, 297, 304, 346.

Continuité (discontinuité): 70, 73-75, 88, 92, 96, 112, 213, 220, 236.

Coréférence (verbo-iconique): 8, 13, 44, 46, 89-90, 106, 313, 340, 351-352, 365-381, 424-425, 458, 463, 471-474, 488.

– simultanée: 369-370, 373-374, 376, 380, 472.

– rétrograde/différée: 374, 376, 380-381, 471.

– intersection référentielle: 370-373, 471.

– disjonction référentielle: 371-372.

– contraste sémantique: 372, 464-465.

Déliation: 8-9, 23-26, 48, 55, 63-64, 142, 195, 336-343, 355, 393, 400-402, 407, 414-419, 426, 460, 485, 493.

Dialogue: 28, 59, 120, 141-142, 160, 169, 185-186, 202-204, 206-211, 239, 282, 293, 374-376, 402, 454, 461-462, 469-470, note 320.

Diègèse, diégétique: 23-25, 31, 61, 73, 104-107, 120, 144-145, 155, 161, 172, 190-193, 201-210, 218, 226-231, 241, 245, 255-257, 260, 278, 288, 304, 335, 337-342, 348-354, 356-368, 371, 395, 407, 415-419, 422, 436-442, 445, 460, 470, 474-478, 489-490, note 639.

Discours intérieur: 8, 12, 108-116, 316, 345-346, 364-365, 382, 425, 431, 440, 475-476, 486, 495.

Dispositif: 21, 31, 33-34, 39-57, 60-61, 66, 78-82, 99, 104-105, 113, 117-118, 130-131, 151, 163, 173, 193, 197-202, 222-224, 243, 250, 263, 341, 386-392, 446, 487.

Distanciation: 22-23, 106-107, 212, 256, 323, 474-478.

Énonciateur, énonciation: 12, 21-22, 29-32, 42, 90, 103, 106-107, 113-116, 169-171, 192-194, 203-204, 208-212, 216, 226, 240, 244, 256, 260-262, 282, 288, 294-295, 317-326, 330, 337, 356-357, 381-447, 452, 468-471, 478, 480-481, 485-487, note 313.

Expérimental (cinéma): 386-387, 395.

Exploitation, salles: 17-18, 24-25, 34-35, 48-49, 53-57, 65-66, 71, 75, 77, 90-91, 95, 102-104, 108-110, 114, 132-133, 136, 139, 150, 157, 221, 247-250, 272-273, 277, 290-294, 304, 310-311, 327, note 86.

Extradiégétique: 24, 89, 113, 144, 177, 193, 209, 236-237, 337-340, 357-360.

Fantasmagorie: 73, 78-81, 137.

Fictionnalité: 326-335, 354, 359, 362, 423, 471, 487.

Fictivité: 423, 425, 428, 487.

Fixation, inscription, enregistrement (des sons): 18-19, 26-27, 80, 88, 91, 95-96, 116-117, 147, 166-167, 194-195, 222-223, 251, 285, 386, 393-395, 401-402, 410-411, 414-415, 479, 485-486, 488, 491.

Focalisation (ocularisation, auricularisation, « point de vue »): 161, 204, 207, 261, 308, 315, 318, 337, 355, 376, 381, 390, 425, 441, 465, 493.

Fonctions (du langage, au sens de Jakobson): 117-129.

Grain (de la voix): 24, 84, 132, 191, 255, 357, 365, 431, 467, 470, 495.

Hétérodiégétique/Homodiégétique: 22, 338-339, 342, 357-358, 361, 364-365, 395, 417, 422, 424, 436, 489.

Instauration (d'un monde): 260, 334, 339, 346-381, 425, 452, 490.

Institutionnalisation, standardisation, MRI: 43-44, 67, 177, 201, 248-249, 269, 280, note 86.

Intégration: 21, 26-27, 56-57, 105, 116, 223-232, 236-239, 276, 345, 428, 491.

Intériorisation, monologue intérieur: 109, 323, 334, 342-343, 395, 425, 440-443, 452, 459-460, 471.

Intermédialité: 197-201, 277-278, 287.

Internet, sites Web: 364, 379, 496-498.

Interrègne: 19, 59, 178, 180, 197-201, 239, 250, 255, 268, 295-296, 303, 491-492.

Intertexte: 74, 158, 298, 426, 429-430, 433, 453.

Juxtadiégétique: 209, 256, 493.

Lanterne (magique): 37, 39, 44, 62-69, note 118, 72-82, 95-97, 99, 156, 158, 220, 243, 326.

Lettriste (cinéma): 386, 395, 419.

Matérialité phonographique: 402, 406, 416-418, 428, 432, 474, 485.

Mégadiégèse: 361-362, 366, 436, 489.

Mise en phase: 407, 415, 485.

Mise en relief: 377-378.

Monde-s possible-s: 31, 353-357, 362-364, 490-491, 498.

Musique: 28, 30, 35, 54, 56, 58-59, 63, 75, 77, 82, 87-89, 92, 102, 108-110, 131-148, 156, 171-172, 211, 218, 221, 227, 229, 236, 241-242, 247-250, 253, 265, 273-274, 278, 297, 302, 306, 309, 360-361, 364, 379, 411, 456-457, 476.

Narrativité, narration, narrativisation: 13, 74, 105, 128, 153-154, 157, 187, 203, 214-218, 224, 230-232, 235-236, 258-260, 270-271, 281-282, 292, 315-322, 335-336, 378, 381, 421, 446.

*Off* (voix, son): 23-26, 64, 150, 257, 342, 344, 416-418, 460.

Oralité: 12, 17-19, 32-38, 51, 72, 75, 78, 80-81, 95, 99-107, 109, 111, 123, 147, 149, 172-181, 184-190, 211, 224, 248-251, 259, 280-282, 285, 288, 296, 318-323, 375, 439, 441-442, 480, 491, 496-497.

Passé: 183, 261, 326-335, 359, 443, 468.

Passion (du Christ): 71-72, 74, note 122, 211-212, 215-216, 374.

Phonographie: 24, 30, 35-36, 39, 41-42, 46-54, 60, 70, 82, 94-95, 107, 109, 222, 264, 286-287, 298, 404-406, 483, 486, 496.

Phonoscènes: 59, 266, 269, 279.

Plan de référence: note 672, 371, 458, 491.

Plaquage: 342, 426, 442, 486.

Pluridiégéticité: 360, 366, 425, 489.

Profilmique: 50, 87, 101, 150, 187, 208, 216, 226, 252, 258, 266, 398.

Prologue: 19, 26, 49, 58, 247, 257, 287-313, 332, 337, 339-340, 342, 352, 360, 425, 450, 452.

Prophonographique: 19, 50, 253, 397, 404-406, 409-410, 481, 484, 488.

Radio-phonie: 16, 42, 168, 175, 177, 180, 185, 234, 257, 277, 396, 420, 432, 485, 489, 496, 498-499.

Redondance: 45, 175, 207, 231-232, 238, 368, 377, 471, 486.

Regard-caméra, regard-public: 302-307, 412, 427.

Rythme: 58-59, 75, 84, 86-87, 89, 92-93, 97, 154, 165, 181, 221, 237, 240-241, 315-316, 419, 456, 473, 476-477, 486.

Série culturelle: 26, note 91, 70, 73, 198, 275, 496.

Synchronisation, synchronisme: 23, 26, 43-44, 48, 55, 57, 81-97, 142, 145, 148, 170-171, 187, 192, 202, 222, 252, 299-300, 368-370, 374, 376-377, 386, 400, 420, 439-440, 444, 456, 471-474.



Théâtre, dimension scénique: 26, 57, 64, 74, 130, 136, 141, note 300, 152, 157, 161, 166, 177, 180-182, 200, 211-212, 216, 219, 227, 232, 238-241, 243, 265-266, 272-274, 278, 281-282, 286-300, 303-304, 329-330, 426-427, 490.

Traduction: 112, 117, 124-129, 153, 166, 168.

Vaudeville (américain): 64, 77, 141, 220-221, 227, 230, 240, 249, 251, 266-267, 272-285, 288-291, 296, 299, 304, 306, 308, 313, 345, 421, 491-492.

Voix-action: 9, 202-213, 223, 226, 233-234, 237, 257, 260, 287, 309, 334, 421, 451, 455, 484, 498.

Voix-attraction: voir attraction.

Voix-explication: 51, 202, 209-210, 335, 349, 360, 484.

Voix-narration: 18, 32, 57, 161, 186, 201-212, 242-260, 303, 315-478, 485-488, 498.

Voix-over: 8, 22-26, 31, 37, 44, 60, 89-90, 107, 115-117, 166-196, 225, 235, 254-255, 259-261, 288, 296, 307-313, 320-324, 326-330, 332-337, 340-346, 352, 356, 360-364, 368-373, 377-380, 387-389, 395, 414, 418, 421-426, 429, 431, 433-435, 437-443, 445, 449-478, 490, 493-498, note 713.



## INDEX DES NOMS PROPRES

- Abel, R. : 221.  
Adam, J.-M. : 128, 225, 331, 336, 350.  
Adler, L. : 473.  
Adorno, Th. et Eisler, H. : 140-143.  
Albera, F. : 32-34, 40, 48, 56, 110-111, 248.  
Alburger, J. R. : 497.  
Allen, J. T. : 281-282, 285.  
Allen, R. C. : 272-273.  
Althusser, L. : 22.  
Altman, R. : 20, 27, 43, 54, 59, 62, 65-66, 72-73, 107, 133, 139, 144-145, 148, 156, 197-200, 221, 228-230, 233-234, 236, 238, 245, 247, 266, 273, 275, 401.  
Anderson, J. L. et Richie, D. : 130.  
Arlaud, R.-M. : 34, 41, 45, 105, 109, 137-138.  
Arletty : 178.  
Arnaud, Cl. : 178.  
Arnheim, R. : 405.  
Arnoldy, E. : 27, 65, 198, 225, 277, 491.  
Arnoux, A. : 146, 265-272.  
Astaïre, Fred : 229, 236-237, 427.  
Aumont, J. : 219.  
Auric, G. : 146-147.  
Aveline, Cl. : 167.
- B -
- Baker, J. : 393.  
Balazs, B. : 439-440.  
Banville, Th. de : 78.  
Barilier, E. : 238.  
Barjavel, R. : 159-160.  
Barnier, M. : 27, 148, 265.  
Barthes, R. : 13, 24, 74, 127, 203, 286-287, 344, 369, 373, 376, 468, 470, 488, 499.  
Bazin, A. : 16, 349, 362.  
Benayoun, R. : 239, 472.  
Benjamin, W. : 187-190.  
Benveniste, E. : 103, 186, 320-322, 383, 467-468.  
Berg, A. : 238.  
Bergel, P. : 136.  
Berkeley, B. : 226-227, 296.  
Birdwhistell, R. : 85.  
Black, A. : 75, 96.  
Blair, L. : 415.  
Bogart, H. : 336-337, 340-342, 426.  
Bogdanovich, P. : 299.  
Bono, F. : 265, 271.  
Booth, W. C. : 319.  
Bordwell, D. : 251-252, 325, 348, 387.  
Bottomore, S. : 129.  
Branigan, E. : 435, 438.  
Brecht, B. : 22, 107, 143, 212, 474, 477-478.  
Brinckmann, Ch. N. : 332, 438-439.  
Browning, T. : 303, 306-307, 310.

Burch, N.: 11, 19, 74, 91, 101, 103, 106-107, 130-131, 178, 215, 248-250, 253, 280, 492.  
 Bush, S.W.: 109-110, 244, 249.  
 Byron: 309-310, 312.

– C –

Cagney, J.: 278, 295.  
 Campan, V.: 461.  
 Canfield, M. C.: 192, 284.  
 Canjels, R.: 288, 290-291.  
 Cantor, E.: 277-278.  
 Canudo, R.: 58.  
 Carbine, M.: 35.  
 Cardinal, S.: 100, 223, 408.  
 Carlier, Ch.: 465.  
 Carol, M.: 255, 262.  
 Carton, P.: 192.  
 Céline, L.-F.: 136.  
 Cendrars, B.: 92-95.  
 Chaplin, Ch. (Charlot): 93, 165-166, 276.  
 Chartier, J.-P.: 323, 439.  
 Chateau, D.: 351-352, 383.  
 Châteauvert, J.: 24, 31-32, 46, 122, 128, 304-305, 334, 346, 356, 389, 395, 421, 423, 436, 429, 431, 436, 441-442.  
 Chevalier, M.: 166, 229.  
 Chion, M.: 25, 28, 60, 83, 140, 172, 297, 417, 420.  
 Chrétien de Troyes: 208.  
 Clair, R.: 12, 166, 210, 234-235, 239.  
 Clark, E. C.: 497.  
 Cocteau, J.: 174, 430, 480.  
 Codi, B.: 50-51.  
 Cohen, K.: 475.  
 Cohen-Seat, G.: 347-348.  
 Cœuroy, A. (et Clarence, G.): 404-406.  
 Coissac, G.-M.: 44-45, 50, 68-72, 77, 82, 91-92, 95, 99, 249.

Colin, M.: 29, 351, 441.  
 Collinet, M.: 266-267.  
 Colpi, H.: 457, 459.  
 Comolli, J.-L.: 249, 413-414.  
 Cosandey, R.: 68.  
 Coustet, E.: 139, 404.  
 Cowl, D.: 242.

– D –

Dalbe, G.: 135-136, 169.  
 Dalle, B.: 478-480.  
 Dällenbach, L.: 411.  
 Daney, S.: 23.  
 Darnell, L.: 396.  
 Dauman, A.: 453.  
 Decremps, H.: 79.  
 Delacommunes, Ch.: 420, note 162.  
 Deleuze, G.: 460, 477.  
 Della Porta, G.: 63.  
 Delluc, L.: 405.  
 Delubac, J.: 194.  
 Devos, R.: 242.  
 Diamant-Berger, H.: 264.  
 Dickson, W. K. L.: 86.  
 Doane, M. A.: 116, 341.  
 Doniol-Valcroze, J.: 179.  
 Doyle, A. C.: 52.  
 Dubois, Ph.: 213-214.  
 Ducrot, O.: 390-391, 393.  
 Duhamel, G.: 174.  
 Dulac, G.: 59, 456.  
 Dumas, A.: 300.  
 Duras, M.: 366, 372, 434, 450-451, 453, 455, 462, 464-468, 473, 477-481, 493.  
 Dureau, G.: 132.  
 Durrer, S.: 185, 210.

– E –

Eco, U.: 84-85, 353-354, 363-364, 391, 488, 491.  
 Edison: 24, 47, 53, 60, 86, 95, 128, 153, 222, 273, 275-276.

Eikhenbaum, B.: 22, 108, 110-114, 316, 324, 382, 475, 486.  
Eisenstein, S. M.: 87, 219, 254, 283, 390.

– F –

Fairbanks, D.: 299-303.  
Fellini, F.: 393.  
Feuer, J.: 232.  
Finley, W.: 410.  
Fischinger, O.: 405.  
Fónagy, I.: 123, 237-238, 463.  
Fonda, H.: 204-205.  
France, A.: 173.  
Frank, N.: note 314, note 598.  
Fréhel: 181.  
Freud, S.: 48, 455, 475.  
Friedeman, K.: 329-330.  
Fuller, L.: 245.  
Fusco, G.: 457.

– G –

Gance, A.: 59, 94, 456.  
Gardies, A.: 74, 349-351, 490.  
Gardner, A.: 337.  
Garland, J.: 234.  
Garroni, E.: 112-114.  
Gaudreault, A.: 21, 33-34, 37-38, 46, 65, 74, 87-88, 104, 106, 149, 153, 198-199, 213-219, 221, 270-271, 279, 281-283, 325, 330, 340, 345, 396, 398, 400, 403, 421, 436, 446, 489, 491.  
Gaumont, L.: 54, 59-60, 247, 266-267, 279, 300.  
Geduld, H. M.: 54.  
Genette, G.: 21, 24, 67, 317-318, 322, 325, 337-340, 343-344, 347-350, 357-358, 361, 366, 437, 446, 489.  
Girardi, A.: 265.  
Giraudoux, J.: 174.  
Godard, J.-L.: 52, 178-179, 190, 204, 344, 493-494.

Goffman, E.: 169, 181, 428.  
Gomery, J.-D.: 279.  
Gorki, M.: 61, 137, note 266, 140.  
Grave, S.: 192.  
Gravey, F.: 178.  
Griffith, D. W.: 58, 218, 279, 346, 489.  
Guido, L.: 58, 476.  
Guitry, L.: 173.  
Guitry, S.: 8, 19, 165-196, 240, 242, 256, 282-283, 329-330, 430, 432, 480.  
Gunning, T.: 11, 42, 80-91, 107, 130, 201, 213-216, 218-221, 223, 227, 246-247, 254, 270, 283, 346, 491.

– H-I-J –

Hamburger, K.: 329-332, 392, 439, 490.  
Hansen, M.: 100.  
Harris, S. H.: 241.  
Hauterives, H. de Grandsaigne de: 95, 135.  
Hays, W.: 266, 296-298.  
Hediger, V.: 288-291, 310-312, note 546.  
Henderson, B.: 322, 329.  
Hepworth, C.: 245.  
Hjelmslev, L.: 369, 397.  
Hoffman, E.T.A.: 52.  
Holm, C.: 338.  
Holmes, B.: 72-73.  
Howe, L.: 107.  
Huygens, C.: 81.  
Huygens, Ch.: 81.  
Isou, I.: 386.  
Jakobson, R.: 21, 117-129, 441.  
Johnson, W.: 266.  
Jolson, Al.: 278, 426.  
Jost, F.: 88-89, 207, 355, 386, 389-391, 393-394, 425, 431, 435, 437-441, 465, 486.  
Jullier, L.: 29-30, 243, 401-402, 406, 420, 476.

## – K –

Kast, P.: 179.  
 Keim, J. A.: 45, 50, 114, 167-168.  
 Keit, A.: 174.  
 Keith et Albee: 272.  
 Kircher, A.: 62-63, 76, 78, 130.  
 Kolb, B., et Wishaw, J.: 379.  
 Komatsu, H. K. et Loden, F.: 134.  
 Kozloff, S.: 31-32, 259, 307-308, 421-424, 433.  
 Kracauer, K.: 138-139, 240, 247, 291.  
 Kuroda, S.-Y.: 320, 397.

## – L –

Labarthe, A.-S.: 476.  
 Labov, W.: 259-260.  
 Lacasse, G.: 33-36, 46, 71, 78, 88, 117, 121, 124, 127-128, 131, 134-135, 181, 246, 248, 284.  
 Laffay, A.: 31, 138, 322-327, 329, 331, 344, 400, 422, 439, 486, 493, note 597, note 600, note 601.  
 Lanchester, E.: 308-309.  
 Lapiere, A.: 76.  
 Lardeau, Y.: 175-176.  
 Lastra, J.: 19, 43, 287, 397.  
 Leavitt, M. B.: 274.  
 Lee, Ch.: 493.  
 Le Fraper, Ch.: 263.  
 Le Guay, Ph.: 169-170.  
 Leloir, M.: 301-302.  
 Lelouch, Cl.: 394.  
 Lemaitre, M.: 366, 386-387.  
 Le Men, S.: 37.  
 Leroy, Cl.: note 165.  
 Lesage, A.-R.: 338.  
 Leskov, N.: 187, 189.  
 Leutrat, J.-L.: 439.  
 Lewis, D.D.: 374.  
 Lindekens, Ch.: 373.  
 Lion, R.: 108-110.  
 Llewellyn, R.: 303, 332-333.  
 Loew, M.: 272-273, 276.

London, K.: 138-141.  
 Lordier, G.: 58, 64, 141.  
 Lotman, I.: 380-381, 458.  
 Lumière, frères: 47, 69, 137, 141, 271, 403.

## – M –

Madis, A.: 173.  
 Malthête-Méliès, M.: 149-163.  
 Mandelstamm, V.: 289, 300, 311-312.  
 Mannoni, L.: 62, 64, 75-76, 78.  
 Marie, M.: 271.  
 Marion, Ph.: 198.  
 Marker, Ch.: 88, 366, 411, 453.  
 Marx, frères: 238-242, 278.  
 Masson, A.: 27, 228.  
 Matlock, J.: 78.  
 Mc Cambridge, M.: 415.  
 Mc Cay: 276.  
 Mc Donald, J.: 229.  
 Méliès, A.: 134-135.  
 Méliès, G.: 90, 126, 149-163, 200, 202, 247, 271, 297.  
 Messter, O.: 265, 271, 279.  
 Metz, Ch.: 25, 32, 84, 88, 118, 120-122, 163, 209-210, 325, 344, 369, 373, 382, 387-389, 397, 413, 416, 441, 487-489.  
 Miller, F.: 495.  
 Milner, J.-Cl.: 370-371.  
 Mitchum, R.: 443.  
 Mitry, J.: 375-377, 471.  
 Moholy-Nagy, L.: 405-406.  
 Monclair, Didier: note 162.  
 Monet, Cl.: 173.  
 Moore, M.: 494.  
 Morand, P.: 174.  
 Moreno, M.: 183.  
 Morin, E.: 10, 143.  
 Morlay, G.: 178.  
 Mottet, J.: 281-283, 345.  
 Mouren, Y.: 457.  
 Müller, J.: 229-232.

Muller, M. : 317-318.  
 Murch, W. : 30.  
 Murcia, Cl. : 450.  
 Musser, Ch. : 119, 128-130, 215-216, 246, 272.

– N-O –

Nedelec, A.-S. : 181, 430-432.  
 Nicéron, J.-F. : 63-64.  
 Niney, F. : 61.  
 Odin, R. : 21, 31, 84, 192, 232, 389, 407, 486, 490.  
 Okada, E. : 449, 452, 455, 469.  
 Ophuls, M. : 19, 254-262.

– P –

Pagnol, M. : 180-181.  
 Pains, D. : 176.  
 Panofsky, E. : 172.  
 Pasolini, P. P. : 85.  
 Pathé, frères : 71, 109, 158, 264, 286.  
 Pavel, Th. : 355.  
 Pellerin, J.-C. : 158.  
 Pendakur, M. : note 420.  
 Percheron, D. : 23, 339.  
 Perrault, Ch. : 158-161.  
 Perriault, J. : 39, 62, 64, 76.  
 Petitjean, A. : 350.  
 Peytard, J. : 318-320.  
 Pfenninger, R. : 405.  
 Piaget, J. : 110-111, 327, 350, note 199.  
 Pidgin, Ch. F. : 76.  
 Pingaud, B. : 450.  
 Piotrovski, A. : 405.  
 Pisano, G. : 27, 286.  
 Portes, J. : 51, 273, 277-278, 285.  
 Pouillon, J. : 315-316.  
 Pozner, V. : 120, 134-135.  
 Prieur, J. : 80, 97, 102.  
 Proust, M. : 96-97, 158, 317-318.

– R –

Rabon, I. : 102-103, 285-286.  
 Ravar, R. : note 783.  
 Raynauld, I. : 71-72, 212.  
 Régnauld, L. : 34, 105.  
 Resnais, A. : 11, 19, 165, 178-180, 234-235, 238, 328, 334, 344, 359, 364-365, 381, 393, 432, 449-481, 492-493.  
 Reusse, A. de : 121.  
 Reynaud, E. : 77.  
 Ricœur, P. : 128, 243, 321-322, 329, 331-332, 490.  
 Riva, E. : 381, 452, 455, 458, 462-463, 466, 469-470, 474, 476-477, 479.  
 Robbe-Grillet, A. : 359, 372, 433-434, 450, 456, 493.  
 Robertson, E. R. : 64, 78-81.  
 Roché, H.-P. : 434.  
 Rogers, J. : 236.  
 Rohmer, E. : 208, 211-212, 451.  
 Ropars-Wuilleumier, M.-Cl. : 450, 475.  
 Rostand, E. : 173.  
 Rouch J. : 16, 187, 414.  
 Russell, L. : 208.  
 Ruttmann, W. : 366.  
 Ryan, M.-L. : 354-355.

– S –

Sadoul, G. : 200.  
 Saint-Saëns, C. : 88, 132.  
 Sanaker, J. K. : 381.  
 Saulsberry, R. : 496, 498.  
 Sawato, M. : 50, 120.  
 Schaeffer, J.-M. : 382-385, 428.  
 Schwarz, A. : 280-281.  
 Scorsese, M. : 343, 374, 494.  
 Shelley, M. : 308-310, 312-313.  
 Shelley, P. : 309.  
 Siclier, J. : 179.  
 Silverman, K. : 260-261.

Sinatra, F.: 234.  
 Sirois-Trahan, J.-P.: 117-118, 279.  
 Skal, D. J.: 306-308.  
 Smith, M.: note 758.  
 Snyder, R. W.: 275, 280, 284.  
 Sokourov, A.: 14, 494.  
 Sopocy, M.: 51, 65-67.  
 Souriau, A.: note 641, note 643.  
 Souriau, E.: 19, 347-349, 353-354, 489.  
 Stocker, B.: 306.  
 Stoddard, J.: 72-73.  
 Straub et Huillet: 10, 204, 366, 414, 457, 480.  
 Suhamy, H.: 368-369.  
 Suwa, N.: 478-481.

– T-U –

Tan, Ed. S.: note 758.  
 Tati, J.: 407.  
 Tedesco, J.: 270.  
 Todorov, T.: 320.  
 Tortajada, M.: 40, 48, 56.  
 Toulet, E.: 138-139.  
 Truffaut, F.: 165, 170, 178-181, 186, 190, 256, 430, 434, 480, 493.  
 Tynianov, I.: 405.  
 Ullmann, Liv: 209.  
 Ustinov, P.: 254-262, 276.

– V –

Vanoye, F.: 345.  
 Van Rossum-Guyon, F.: 319.  
 Van Sloan, Ed.: 304-306.  
 Varda, A.: 450, 493.  
 Vernet, M.: 426-427.  
 Vittorini, E.: 455, 457, 480.  
 Vogeeler, S.: 355-356.  
 Vuillermoz, E.: 136-140.  
 Vygotsky, L.: 110-111, 327.

– W-Y-Z –

Walgensten, Th.: 63.  
 Waxman, F.: 309.  
 Weinrich, H.: 329-331, 378, 490.  
 Welles, O.: 10, 177, 373, 394, 441, 446, 494.  
 White, S.: 254.  
 Williams, A.: 43, 401.  
 Williamson, J.-A.: 65-66.  
 Wiseman, F.: 207.  
 Wolfe, Ch.: 115, 237, 277.  
 Wong Kar-Wai: 494.  
 Wyss, A.: 237.  
 Yabus, A.L.: 379.  
 Zukor, A.: 276.  
 Zumthor, P.: 12, 16, 20, 26, 32-33, 35-36, 100-103, 106-107, 116, 123, 131-132, 181, 184, 203, 225, 233, 235, 246, 252-253, 281-282, 352, 409, 485, 492, 496, 498.



# INDEX DES TITRES DE FILMS

(ORDRE ALPHABÉTIQUE DES TITRES FRANÇAIS)

- The Adventures of Dollie*, D. W. Griffith, 1908: 279.
- Les Affranchis (Goodfellas)*, Martin Scorsese, 1990: note 826.
- Les Affiches en goguette*, Georges Méliès, 1905: 151.
- L'Âge de l'innocence (The Age of Innocence)*, Martin Scorsese, 1993: note 826.
- L'Aigle des mers (The Sea Hawk)*, Frank Lloyd, 1924: 295.
- Alien, le huitième passager (Alien)*, Ridley Scott, 1979: 419.
- Aliens*, James Cameron, 1986: 417-419.
- Alien 3*, David Fincher, 1992: 418-419.
- Alien 4: La Résurrection (Alien: Resurrection)*, Jean-Pierre Jeunet, 1997: 419.
- L'Anglaise et le duc*, Eric Rohmer, 2002: 208.
- L'Année dernière à Marienbad*, Alain Resnais, 1961: 359, 432-433, 474.
- À propos de Nice*, Jean Vigo, 1930: 189.
- L'Arroseur arrosé*, frères Lumière, 1895: 217.
- L'Assassinat du Duc de Guise*, André Calmettes et Charles Le Bargy, 1908: 88, 132.
- Assassins et voleurs*, Sacha Guitry, 1956: 242.
- Assurance sur la mort (Double Indemnity)*, Billy Wilder, 1944: 324.
- L'Atlantide (Die Herrin von Atlantis)*, G.-W. Pabst, 1932: 270.
- À tombeau ouvert (Bringing out the Dead)*, Martin Scorsese, 1999: note 826.
- Attack on a China Mission*, James Williamson, 1900: 51.
- Avalon*, Mamuro Oshii, 2001: 360-364.
- L'Aveu*, Costa-Gavras, 1970: 414.
- Le Ballet mécanique*, Fernand Léger et Dudley Murphy, 1924: 93-94.
- Barbe-Bleue*, Georges Méliès, 1901: 151, 154-155, 158-163.
- Berlin Express*, Jacques Tourneur, 1949: note 320.

- Blade Runner*, Ridley Scott, 1982: 346.
- Bonne chance*, Sacha Guitry, 1935: 194-195.
- Boarding School Girls*, Edwin S. Porter, 1905: 128.
- Bob Hampton of Placer*, Marshall Neilan, 1921: 292.
- Le Brave soldat Svejk*, Jiri Trnka, 1955: note 315.
- Brigadoon*, Vincente Minnelli, 1954: 235.
- Butterfly Murders*, Tsui Hark, 1979: 343.
- Cabinet du Dr Caligari (Das Cabinet des Dr Caligari)*, Robert Wiene, 1919: 293-294.
- Le Cake Walk infernal*, Georges Méliès, 1903: 152.
- Cars*, John Lasseter, 2006: 490.
- Les Cartes vivantes*, Georges Méliès, 1904: 152.
- Casablanca*, Michael Curtiz, 1942: 227.
- Casino*, Martin Scorsese, 1995: 343, 494.
- Cavalcade*, Frank Lloyd, 1933: 375.
- C'est arrivé près de chez vous*, Remy Belvaux, 1992: 412-414.
- Ceux de chez nous*, Sacha Guitry, 1915: 173, 176-178, 180.
- Chaînes conjugales (A Letter to Three Wives)*, Joseph L. Mankiewicz, 1949: 337-338, 395-396, 437.
- Chantage (Blackmail)*, Alfred Hitchcock, 1929: 302.
- Chant de la fidèle Chunhyang*, Im Kwon Taek, 2000: 235.
- Le Chant de la fleur écarlate*, Teuvo Tulio, 1938: 148.
- Le Chant du Styrene*, Alain Resnais, 1958: 475.
- Le Chanteur de jazz (The Jazz Singer)*, Alan Crosland, 1927: 231, 269, 277-278, 426.
- Le Chanteur du dessus (Top Hat)*, Mark Sandrich, 1935: 236.
- Charlie et la chocolaterie (Charlie and the Chocolate Factory)*, Tim Burton, 2005: 492-493.
- Le Chaudron infernal*, Georges Méliès, 1903: 152.
- Chicago*, Bob Marshall, 2002: 492.
- Citizen Kane*, Orson Welles, 1941: 260, 324, 373.
- Cléo de 5 à 7*, Agnès Varda, 1961: 459.
- Cléopâtre (Cleopatra)*, Joseph L. Mankiewicz, 1963: note 620.
- La Comtesse aux pieds nus (The Barefoot Contessa)*, Joseph L. Mankiewicz, 1954: 335-337, 340-343, 345, 426.
- La Couleur de la grenade*, Sergej Paradjanov, 1968: 235.
- Creature Comforts*, Nick Park, 1989: 495.
- De Jeanne d'Arc à Philippe Pétain*, Sacha Guitry, 1942: 174-178.

- Démolition d'un mur*, frères Lumière, 1895: 47.
- Désiré*, Sacha Guitry, 1937: note 369.
- Detour*, Edgar G. Ulmer, 1946: 442.
- Les Deux Anglaises et le continent*, François Truffaut, 1971: 178, 381, 430.
- 2001, l'Odyssée de l'espace (2001: A Space Odyssey)*, Stanley Kubrick, 1968: 417.
- Discussion de M. Jansen et de M. Lagrange*, frères Lumière, 1895: 141.
- Les Dix Commandements (The Ten Commandments)*, Cecil B. DeMille, 1923: 294-295.
- Don Juan*, Alan Crosland, 1926: 297.
- Doom*, Andrej Bartowiak, 2005: note 620.
- Dracula*, Tod Browning, 1931: 303, 306-307, 309-310.
- Duel*, Steven Spielberg, 1975: 342-343.
- Du silence et des ombres (To Kill a Mockingbird)*, Robert Mulligan, 1962: 419.
- L'Enfant sauvage*, François Truffaut, 1970: 430.
- Les Enfants du Paradis*, Marcel Carné, 1944: 286.
- L'Étrange Noël de M. Jack (Nightmare before Christmas)*, Henry Selick, 1993: 493.
- Ève (All about Eve)*, Joseph L. Mankiewicz, 1950: 167.
- Et vogue le navire (E la nave va)*, Federico Fellini, 1983: 393.
- L'Exorciste (The Exorcist)*, William Friedkin, 1973: 414-416.
- L'Explorateur en folie (Animal Crackers)*, Viktor Heerman, 1930: 239.
- Le Fabuleux destin d'Amélie Poulain*, Jean-Pierre Jeunet, 2001: 494.
- Faisons un rêve*, Sacha Guitry, 1936: note 369.
- Le Fantôme du Paradis (Phantom of the Paradise)*, Brian de Palma, 1974: 410.
- Le Faux coupable (The Wrong Man)*, Alfred Hitchcock, 1956: 342.
- Les Feux de la rampe (Limelight)*, Charlie Chaplin, 1952: 276.
- Les Fiançailles rouges*, Roger Lion, 1926: 108.
- La Fiancée de Frankenstein (The Bride of Frankenstein)*, James Whale, 1935: 303, 307-315.
- Final Fantasy: Les Créatures de l'esprit (Final Fantasy: The Spirits Within)*, Hironobu Sakaguchi, 2001: 402.
- La Flèche brisée (Broken Arrow)*, Delmer Daves, 1950: note 620.
- Frankenstein*, James Whale, 1931: 303-312.
- Gangs of New York*, Martin Scorsese, 2002: 374-375, note 826.
- Geneviève de Brabant*, Pathé, 1907: 158.
- Gertie le dinosaure*, Winsor McCay, 1914: 276.

- Ghost in the Shell*, Mamuro Oshii, 1995: 362.
- Les Grandes manœuvres*, René Clair, 1955: 210.
- La Griffé du passé (Out of the Past)*, Jacques Tourneur, 1947: 443-444.
- Guernica*, Alain Resnais, 1950: 453.
- La Guerre des mondes (War of the Worlds)*, Steven Spielberg, 2005: 490.
- La Guerre est finie*, Alain Resnais, 1966: 334.
- Hallelujah*, King Vidor, 1929: 233-234.
- Hermann Slobbe, enfant aveugle (Blind Kid II)*, Johan Van der Keuken, 1966: 409-410.
- Hidalgo*, Joe Johnston, 2003: 50.
- Hiroshima*, Hideao Sekigawa, 1953: 453.
- Hiroshima mon amour*, Alain Resnais, 1959: 22, 178-179, 342, 449-481.
- Un Homme de tête*, Georges Méliès, 1898: 152.
- L'Homme qui ment*, Alain Robbe-Grillet, 1968: 10, 372.
- How it Feels to Be Run over*, Cecil M. Hepworth, 1900: 390.
- 8 Femmes*, François Ozon, 2002: 238.
- H Story*, Nobuhiro Suwa, 2001: 478-481.
- Ils étaient des hommes (The Men)*, Fred Zinneman, 1950: note 620.
- L'Incroyable destin d'Harold Crick (Stranger than Fiction)*, Marc Forster, 2006: 24.
- Les Indestructibles (The Incredibles)*, Brad Bird, 2004: 495.
- Les Infiltrés (The Departed)*, Martin Scorsese, 2006: note 826.
- Les Innocents de Paris (Innocents of Paris)*, Richard Wallace, 1929: 269.
- I Want to Go Home*, Alain Resnais, 1989: 364, 449.
- Jeanne et le garçon formidable*, Olivier Ducastel et Jacques Martineau, 1998: 238.
- La Jetée*, Chris Marker, 1962: 88.
- Jeux dangereux (To Be or not To Be)*, Ernst Lubitsch, 1942: note 620.
- Le Jour se lève*, Marcel Carné, 1939: 457.
- Jules et Jim*, François Truffaut, 1962: 178, 345, 434.
- Kill Bill*, vol. 1, Quentin Tarantino, 2003: 363, 494-495.
- King Kong*, Merian Cooper et Ernest Schoedsack, 1933: 363.
- The Lady*, Frank Borzage, 1925: 293.
- La Lectrice*, Michel Deville, 1988: 421-422.
- La Légende de Rip van Vinckle*, Georges Méliès, 1905: 90.
- Lettre de Sibérie*, Chris Marker, 1958: 367.
- La Ligue des Gentlemen extraordinaires (The League of Extraordinary Gentlemen)*, Stephen Norrington, 2003: 407.

- Lola Montès*, Max Ophuls, 1955: 19, 254-262, 276, 408.
- The Lonedale Operator*, D. W. Griffith, 1911: 279.
- Un long dimanche de fiançailles*, Jean-Pierre Jeunet, 2004: note 320.
- Les Looney Tunes passent à l'action (Looney Tunes: Back in Action)*, Joe Dante, 2003: 364.
- Madame Bovary*, Claude Chabrol, 1991: 434.
- Le Magicien d'Oz (The Wizard of Oz)*, Victor Fleming, 1939: 492.
- La Maison de la 92<sup>e</sup> Rue (House on the 92th Street)*, Henry Hathaway, 1945: note 620.
- Le Masque de fer (The Iron Mask)*, Allan Dwan, 1929: 298-301.
- Matrix (The Matrix)*, Andy et Larry Wachowski, 1999: 362.
- Le Mélomane*, Georges Méliès, 1903: 152.
- Le Mépris*, Jean-Luc Godard, 1963: 190.
- Le Million*, René Clair, 1931: 166.
- Moi un noir*, Jean Rouch, 1958: 187.
- M. Arkadin (Confidential Report)*, Orson Welles, 1955: 441.
- Laura*, Otto Preminger, 1944: note 313, 396.
- Mon oncle*, Jacques Tati, 1958: 407.
- Mon oncle d'Amérique*, Alain Resnais, 1980: 179.
- Monty Python, Sacré Graal (Monty Python and the Holy Grail)*, Monty Python, 1974: 403-404.
- La Mort aux trousses (North by Northwest)*, Alfred Hitchcock, 1959: 207.
- Le Mot de Cambronne*, Sacha Guitry, 1936: note 369.
- Moulin Rouge*, Baz Luhrman, 2001: 492.
- Muriel*, Alain Resnais, 1963: 449.
- Le Mystère de la chambre jaune*, Marcel L'Herbier, 1930: 193.
- Les Noces funèbres (Corpse Bride)*, Tim Burton et Mike Johnson, 2005: 493.
- Noix de coco (Cocoanuts)*, Robert Florey et Joseph Santley, 1929: 239-242.
- Nouvelles luttes extravagantes*, Georges Méliès, 1900: 152.
- Nuit et brouillard*, Alain Resnais, 1955: 328, 453, 463-464.
- Ombres blanches (White Shadows in the South Seas)*, W. S. Van Dyke et Robert Flaherty, 1928: 145-147.
- On connaît la chanson*, Alain Resnais, 1997: 234-235, 393, 449, 474.
- L'Opéra de quat'sous (Die Dreigroschenoper)*, Georg Wilhelm Pabst, 1930: 138, 255.
- Orochi*, Buntaro Futagawa, 1925: 50.
- Othon*, Jean-Marie Straub et Danièle Huillet, 1969: 413-414.
- Ouvert pour cause d'inventaire*, Alain Resnais, 1946: 456.

- Ouvriers, paysans (Operai, contadini)*, Jean-Marie Straub et Danièle Huillet, 2001 : 366.
- Partie de campagne*, Jean Renoir, 1936 : 192.
- Passage à l'acte*, Martin Arnold, 1993 : 419.
- La Passion de Jeanne d'Arc*, Carl Theodor Dreyer, 1928 : 373.
- Pas sur la bouche*, Alain Resnais, 2003 : 449.
- Perceval le Gallois*, Eric Rohmer, 1978 : 208, 212.
- Les Perles de la couronne*, Sacha Guitry, 1937 : note 369.
- Pierrot le fou*, Jean-Luc Godard, 1965 : 242.
- Le Plaisir*, Max Ophuls, 1951 : 442.
- Playtime*, Jacques Tati, 1967 : 407.
- Porte des Lilas*, René Clair, 1957 : 210.
- Premiers pas de bébé*, frères Lumière, 1895 : 217.
- Prologues (Footlight Parade)*, Lloyd Bacon, 1933 : 295.
- Qu'elle était verte ma vallée (How Green Was my Valley)*, John Ford, 1941 : 326-327, 332-336.
- Qui veut la peau de Roger Rabbit? (Who Framed Roger Rabbit?)*, Robert Zemeckis, 1988 : 363.
- Rashômon*, Akira Kurosawa, 1950 : 356.
- Rebecca*, Alfred Hitchcock, 1940 : 312, note 620.
- Repas de bébé*, frères Lumière, 1895 : 217.
- Retour vers le futur (Back to the Future)*, Robert Zemeckis, 1985-1990 : 359.
- Roger and Me*, Michael Moore, 1989 : 225.
- Le Roman d'un tricheur*, Sacha Guitry, 1936 : 8, 19, 165-196, 283, 312, 330, 343, 430, 432, 490.
- La Ronde*, Max Ophuls, 1950 : 257.
- La Roue*, Abel Gance, 1923 : 59.
- La Ruée vers l'or (Gold Rush)*, Charlie Chaplin, 1925-1942 : 166.
- Sans soleil*, Chris Marker, 1983 : 411.
- Le Secret derrière la porte (The Secret Beyond the Door)*, Fritz Lang, 1948 : note 313.
- Le Seigneur des anneaux: La Communauté de l'anneau (The Lord of the Rings: The Fellowship of the Ring)*, Peter Jackson, 2001 : 490.
- Sicilia!* Straub et Huillet, 1998 : 457.
- Le Silence est d'or*, René Clair, 1947 : 167.
- Sin City*, Frank Miller et Roberto Rodriguez, 2005 : 495.
- The Sins of Lola Montès*, Max Ophuls, 1955 : 255.
- Si Paris nous était conté...*, Sacha Guitry, 1956 : 188-189.
- Sonate d'automne*, Ingmar Bergman, 1978 : 209.

- Son nom de Venise dans Calcutta désert*, Marguerite Duras, 1976: 10, 366, 372.
- Smoking/No Smoking*, Alain Resnais, 1993: 480.
- Spider-Man*, Sam Raimi, 2002: 490.
- La Splendeur des Amberson (The Magnificent Ambersons)*, Orson Welles, 1942: 177, 260.
- Star Wars*, George Lucas, 1977-2005: note 24, 363.
- Strawberry Blonde*, Raoul Walsh, 1941: 221.
- Sur la piste des Mohawks (Drums along the Mohawks)*, John Ford, 1939: 204-205.
- Les Temps modernes (Modern Times)*, Charlie Chaplin, 1936: 165.
- Terminator*, James Cameron, 1984: 360.
- Terminator II: Le Jugement dernier (Terminator II: Judgment Day)*, James Cameron, 1991: 360.
- Le Territoire des morts (Land of the Dead)*, George A. Romero, 2005: 420.
- The Terror*, Roy del Ruth, 1928: 303.
- Le Thaumaturge chinois*, Georges Méliès, 1904: 152.
- Tous en scène (The Band Wagon)*, Vincente Minnelli, 1953: note 620.
- Tout l'or du ciel (Pennies from Heaven)*, Herbert Ross, 1981: 235.
- Traité de bave et d'éternité*, Isidore Isou, 1951: 386.
- Le Tripot clandestin*, Georges Méliès, 1905: 152.
- Troie (Troy)*, Wolfgang Petersen, 2004: 490.
- Tron*, Steven Lisberger, 1982: 363.
- Une auberge à Tokyo*, Yasujiro Ozu, 1935: 176.
- Une nuit à l'opéra (A Night at the Opera)*, Sam Wood, 1935: 239.
- Une nuit en enfer (From Dusk till Dawn)*, Robert Rodriguez, 1996: 363.
- Une sale histoire*, Jean Eustache, 1977: 204.
- Une séance Méliès*, Jacques Mény, 2001: 149-163.
- Unsere Afrikareise*, Peter Kubelka, 1961-1966: 414.
- Vaudou (I Walked with a Zombie)*, Jacques Tourneur, 1947: note 620.
- Videodrome*, David Cronenberg, 1982: 363.
- La Vie est un roman*, Alain Resnais, 1983: 449.
- Vitaphone Shorts*, Warner, 1926-1927: 266, 269, 296-297, 299, 491.
- Voice: off*, Donald Cumming, 2003: 410.
- Le Voyage dans la lune*, Georges Méliès, 1902: 155-157.

# TABLE DES MATIÈRES

<b>PREFACE DE FRANCOIS ALBERA</b> .....	7
<b>PROLOGUE</b> .....	17
La voix et l'image, ou l'humain dans la machine .....	17
Axes méthodologiques .....	19
Voix <i>in, off, over</i> : questions de déliaison .....	23
Énonciation et voix-over .....	30
Le bonimenteur du cinéma des premiers temps, une pratique orale .....	32
<b>I. VOIX VIVE ET DISPOSITIFS (PRE)CINEMATOGRAPHIQUES</b> .....	39
I.1 Préalables théoriques: voix et locuteur dans les dispositifs audiovisuels .....	39
I.2 Conférencier, bonimenteur et «phantasmagore»: le «cinéma» <i>parlé</i> .....	56
I.3 Voix vive et synchronisation .....	81
<b>II. LA MEDIATION BONIMENTORIELLE</b> .....	99
II.1 Le boniment comme performance .....	99
II.2 Boniment et «discours intérieur» .....	108
II.3 Les fonctions du boniment .....	117
II.4 La médiation spectateur ↔ machinerie: l'«idéologie» des discours sur le bruit du projecteur .....	130
II.5 <i>Une séance Méliès</i> bonimentée par Madeleine Malthête-Méliès .....	149



<b>III.</b>	<b>LE « BONIMENT » OVER DU ROMAN D'UN TRICHEUR</b> . . . . .	165
III.1	Un usage singulier de la voix . . . . .	165
III.2	Guitry et l'oralité: d'une certaine « modernité » cinématographique . . . . .	180
III.3	Le <i>théâtre</i> de l'exhibition vocale . . . . .	108
III.4	Mémoires et roman . . . . .	185
III.5	Guitry, « conteur » au sens de Walter Benjamin . . . . .	186
III.6	La réflexivité complice d'un générique parlé . . . . .	190
<b>IV.</b>	<b>LES FONDEMENTS THEORIQUES</b>	
	<b>DE LA VOIX-ATTRACTIONS</b> . . . . .	197
IV.1	L'intermédialité de l'interregne . . . . .	197
IV.2	Les régimes vocaux du cinéma . . . . .	201
IV.3	Attraction/intégration (narrative): une conception graduelle . . . . .	213
IV.4	Perpétuation des attractions sous forme vocale . . . . .	227
IV.5	L'intégration musicale de la parole . . . . .	236
IV.6	La voix-attraction parlée: l'exemple des Marx Brothers . . . . .	238
IV.7	Synthèse des traits définitoires de la voix-attraction et de la voix-narration . . . . .	242
IV.8	Voix-attraction et voix-narration dans <i>Lola Montès</i> . . . . .	254
<b>V.</b>	<b>LA VOIX-ATTRACTION DANS L'HISTOIRE DU CINEMA</b> . . . . .	263
V.1	Les sons comme « attractions » dans le cinéma des premiers temps . . . . .	263
V.2	L'attraction sonore à la fin des années 20: le constat d'Alexandre Arnoux . . . . .	265
V.3	L'attraction vocale de type « vaudeville » . . . . .	272
V.4	L'esthétique du vaudeville dans les prologues du « cinéma » parlé . . . . .	287
V.5	Les prologues des premiers parlants: la scène à l'écran . . . . .	296
V.6	Du monstre à sa fiancée: « création » frankensteinienne de la voix-narration . . . . .	307

<b>VI. VOIX-NARRATION ET ENONCIATION FILMIQUE</b> .....	315
VI.1 La notion de «voix narrative» dans les études littéraires .....	315
VI.2 La «révélation» d'Albert Laffay .....	322
VI.3 La fiction associée au passé de la voix (narrative) .....	326
VI.4 «Corps» du film, corps du locuteur .....	335
VI.5 Voix-over et image: deux instruments pour <i>construire un monde</i> .....	346
VI.6 La voix-over parmi les «mécanismes» de l'énonciation filmique .....	381
<b>VII. LA VOIX-OVER DANS HIROSHIMA MON AMOUR</b> .....	449
VII.1 La voix (voie?) d'une certaine modernité .....	449
VII.2 Trois régimes <i>over</i> : la construction progressive d'un sujet (parlant) fictionnel .....	452
VII.3 Le prologue: fusion des corps, fission audiovisuelle .....	460
VII.4 Les possibilités du dialogue .....	469
VII.5 Synchronisation, coréférence .....	471
VII.6 L'immersion malgré tout .....	474
VII.7 La voix (ré)incarquée: le « <i>remake</i> » de Nobuhiro Suwa .....	478
<b>ÉPILOGUE-S</b> .....	483
Voix-narration et énonciation .....	485
Le monde du film .....	488
La postérité des attractions chantées .....	491
La voix-over aujourd'hui .....	493
Internet: le nouvel essor pour la voix-over .....	496
<b>BIBLIOGRAPHIE</b> .....	501
<b>INDEX</b>	
Index notionnel .....	517
Index des noms propres .....	523
Index des titres de films .....	529