

# UNIVERSITE DE LIMOGES

ECOLE DOCTORALE n° 527 :

« Cognition, Comportement, Langage(s) »

Centre de Recherches Sémiotiques (CeReS)

Thèse

pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ DE LIMOGES

Sciences du langage, spécialité sémiotique

Présentée et soutenue par

Juan Alberto CONDE

Le 16/12/2016

**L'audiovision dans le cinéma d'animation expérimental**

Approche sémiotique

Thèse dirigée par Jacques FONTANILLE

JURY :

Rapporteurs

M. François Jost, Professeur émérite de l'Université de la Sorbonne Nouvelle (Paris 3),

Directeur Honoraire du CEISME

M. Laurent Jullier, Professeur à l'Université de Lorraine, IECA, (Nancy)

Examineurs

M. Anne Beyaert, Professeur des Universités Université Montaigne à Bordeaux, Directrice du MICA, MSHA

M. Jacques Fontanille, Professeur émérite, Université de Limoges, CeReS

M. Sylvie Périneau, Maître de Conférences, Université de Limoges, EHIC

*A Tary, qui m'a appris à écouter les voix des animaux*

# Remerciements

Je tiens tout d'abord à remercier le directeur de cette thèse, M. Jacques Fontanille, pour sa patience et sa diligence dans la correction de mes textes, mais aussi pour ses contributions à la sémiotique et à l'étude du langage sans lesquels ce travail n'existerait pas.

Mes remerciements vont également à M. Álvaro Corral, directeur du département d'humanités à l'Université Jorge Tadeo Lozano, pour avoir géré le soutien institutionnel qui m'a permis de voyager en France et mener à bien mes études doctorales.

Je tiens également à remercier à mes collègues de l'université Jorge Tadeo Lozano : les professeurs Douglas Niño et César Díaz, Carlos Pérez et Diana Guzmán. Nos conversations et discussions permanentes sur les différentes approches de la sémiotique et la passion partagée pour le cinéma ont nourri mon parcours de recherche et mes connaissances sur les différents aspects du cinéma.

À mon ami Diego Sosa et mon frère Camilo Conde, avec lesquels j'ai débuté dans l'étude de l'animation, mes copains dans la route vers les lignes et figures en mouvement, mes plus sincères remerciements.

À mes parents, mes premiers maîtres, dans la connaissance et dans la vie : à ma mère pour le cœur et à mon père, pour la tête. Et aux deux pour l'affection. À ma sœur, qui m'a accueilli en Europe et avec qui j'ai partagé beaucoup de rires et de moments heureux dans le cœur froid du Nord.

À la famille Couteau, ma famille en France, mes amis, les plus sincères remerciements. Et à Marina, l'autre membre colombienne de cette famille, merci d'ouvrir cette porte.

À mes collègues du doctorat à Limoges : tant ceux qui ont réussi comme ceux qui sont partis : Soon Hong, Jia Ding, Ludovic Chatenet, Maria Barreto, Lina Liñán, Stephanie Delugeard, Valeria De Luca et toutes les autres... je tiens à les remercier pour les moments partagés et le soutien mutuel.

Et me remerciement le plus grand à Tary, le phare qui a guidé mes retours, pour me faire savoir qu'elle est toujours avec moi, même à la distance. Et à Emma, Abril et Julio Mayo, nous sommes la tribu.

## Droits d'auteurs



Cette création est mise à disposition selon le Contrat : « **Paternité-Pas d'Utilisation Commerciale-Pas de modification 3.0 France** » disponible en ligne :  
<http://creativecommons.org/licenses/by-nc-nd/3.0/fr/>

# Sommaire

Introduction .....	7
Chapitre 1. Le son au cinéma : des études filmiques à la narratologie audiovisuelle .....	11
Chapitre 2. Le rôle du son dans les études sur l'animation .....	44
Chapitre 3. L'énonciation, le corps et les dispositifs audiovisuels .....	83
Chapitre 4. Dispositifs audiovisuels et cinéma d'animation .....	130
Chapitre 5. Les sons comme événements .....	189
Conclusion .....	238
Bibliographie .....	256
Filmographie .....	264
Annexes .....	272

# Introduction

*Le film découvrira notre environnement acoustique. La voix des choses, le langage intime de la nature. Tout ce qui a son mot à dire en dehors du dialogue humain, qui nous parle dans la grande conversation de la vie et influe continuellement et profondément sur notre pensée et nos sentiments. Du mugissement des vagues, du vacarme de l'usine jusqu'à la mélodie monotone de la pluie d'automne frappant les vitres, jusqu'aux grincements du plancher dans la chambre déserte. Des poètes lyriques ont souvent décrit ces voix lourdes de sens qui nous accompagnent. Le film sonore les représentera, il les fera retentir de nouveau. La sensibilité du micro accroîtra la nôtre.*

**Béla Balázs. *L'esprit du cinéma***

Il est inévitable de commencer chez Disney. Ou peut-être mieux, chez Ub Iwerks et Carl Stalling ? Ou chez Power, propriétaire du système Cinephone<sup>1</sup> ? Ou tout simplement, dans un cimetière à la campagne... Le fond noir du début est la pupille d'un hibou qui hulule dans la nuit. Le vent hulule aussi et... Mais, sont-ils des sons différents ? Ou est-ce une flûte, ou un violon d'orchestre qui imite tant le hibou que le vent ? Ou un sifflet ? On continue. Une horloge fait son tic tac, une cloche tinte, les chauves-souris battent ses ailes (et on les écoute), un chien hurle, des chats miaulent, un squelette sort d'une tombe et claque ses dents et, finalement, la danse commence. On a écouté la voix des choses, la voix des animaux, la voix du vent, et, maintenant, la voix des os, qui deviennent des marimbas. Les squelettes sont musiciens et instruments en même temps. L'un d'eux prend un chat et transforme sa queue en violon. Les corps des squelettes sont démontés et remontés, se transforment et sonnent. Toujours ils sonnent. Sur le fond d'un orchestre qui ne cesse de jouer, mais qui se mêle avec les bruits. Or on n'a pas écouté la voix des choses, mais *la musique des choses*. La fin de la nuit s'approche, les squelettes courent et sont fusionnés en une créature cinétique unique qui à son tour se transforme en une courbe pure, presque abstraite, pour sauter dans la tombe et dans le silence.

Cela est le sujet de ce travail : le rôle du son dans le cinéma d'animation. Bien sûr, chaque substantif de ce titre pourrait être pluriel, car le terme *animation* couvre une multiplicité de techniques, de genres, de styles et de possibilités expressives, et on peut dire la

---

<sup>1</sup> Ou même chez Edvard Grieg, comme nous a fait savoir Michel Chion (1985).

même chose sur le son. Comme on va voir pour commencer, la tradition théorique des études filmiques nous a appris à distinguer, dans la matérialité sonore du cinéma, les bruits, les paroles, les musiques. Mais le cinéma lui-même nous a appris que leurs frontières ne sont pas toujours claires, ni leurs fonctions. Ni les instances responsables de leur existence. Comme on a vu dans le premier exemple, le court métrage *Skeleton Dance* (1929), qui fait partie des *Silly Symphonies* du Studio Disney, la richesse de sa sonorité et de la synchronisation avec les images sont le résultat de plusieurs « voix », y compris même de l'intervention des responsables de la technologie *Cinephone*, un système développé par l'ingénieur William Garrity<sup>2</sup> pour la compagnie de Patrick Powers, l'entrepreneur qui a vendu cette technologie à Disney pour son premier film animé avec du son synchronisé, le célèbre *Steamboat Willie* (1928).

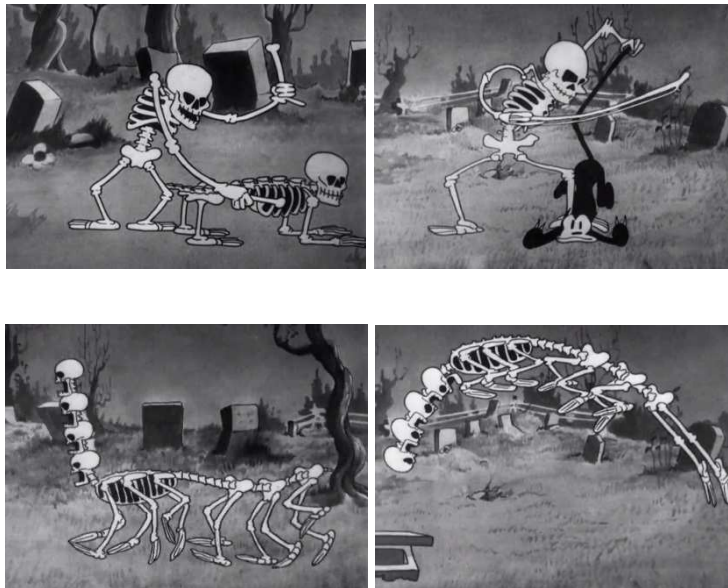


Figure 1 : Photogrammes du film *Silly Symphony: The Skeleton Dance*

Source : [DISNEY, 1929]

Comme ce travail est une thèse de sémiotique, l'inclusion des techniciens et même de la technologie dans l'instance d'énonciation d'une production animée peut sembler excessive, sinon impertinente. Mais pas complètement, si l'on pense que sans le système *Cinephone* ces premiers films synchronisés de Disney, véritables symphonies audiovisuelles, n'existeraient

---

<sup>2</sup> Il ne faut pas oublier que Garrity avait travaillé pour Lee De Forest, l'inventeur du système *Phonofilm*, qu'il a plagié. Voir note 158 du chapitre 4.



pas. Or il est nécessaire, bien sûr, qu'une volonté créatrice prenne la technologie et la mette au service d'une pratique de sens.

En outre, dans mon titre je parle d'*audiovision*, ce qui nous renvoie aussi à l'instance de l'énonciataire. Ce concept, déjà célèbre grâce à son créateur, le théoricien Michel Chion, exprime une activité ou une modalité de la perception qui s'active –est qui est la face subjective– des produits audiovisuels qui, dans le cas de l'animation donne lieu à des expériences singulières. De ce côté, il est vrai que le spectateur ne doit pas connaître les particularités techniques de la synchronisation son/image pour percevoir et s'amuser avec la danse des squelettes, les sons rythmiques et musicaux de leurs os et leurs transformations impossibles. Le lieu de rencontre entre ces instances est donc la scène sémiotique d'une expérience de sens où une forme de textualité (résultat d'une *écriture* sonore et visuelle) inscrite dans un support matériel, s'active et rencontre un corps vivant, qui la transforme en une diégèse, une histoire, un monde ; ou bien aussi en une pure expérience corporelle, rythmique ou vibratoire, somatique, qui est aussi une forme de sens...

Pour aborder cette scène si complexe, j'ai choisi de mettre l'accent sur un aspect particulier : le son, l'écoute, et les approches sémiotiques que se sont occupées de cette dimension de l'expérience perceptive. En ce sens, j'ai voulu profiter de l'expérience du Centre de recherches sémiotiques (CeReS), de l'école doctorale Cognition, comportements, langages, de l'université de Limoges, et de sa ligne d'investigation sur la sémiotique du son. Bien que dans le cinéma et l'audiovisuel, comme dans la vie, l'expérience perceptive fonctionne comme un ensemble, une totalité, l'accent –et parfois, l'« isolement » méthodologique, une sorte de réduction phénoménologique– du son peut nous offrir des aspects inédits sur cette facette de l'expérience que nous procure l'écoute.

Mais ce travail n'est pas une enquête purement théorique. Au contraire, la principale motivation de ma recherche est une passion, un gout, une perplexité : la fascination pour des os qui dansent, des formes et couleurs qui sonnent, des lignes abstraites qui nous font vibrer, rire, bouger. Bref, l'univers expressif du dessin animé.

Cependant, je ne vais pas m'occuper des courts métrages de Disney ni du « cartoon » américain en général, comme on pourrait le penser après l'exemple des squelettes qui m'a aidé comme prétexte pour commencer. Dans ma révision de la littérature existante sur le son dans l'animation, ces productions seront souvent mentionnées, mais, dans mes propres

analyses (chapitres 3 et 4) je vais me concentrer sur un autre type de productions animées que, selon le critère, on peut appeler « indépendantes », « expérimentales », ou même « abstraites ». Néanmoins, pour éloignées que semblent ces productions de l'animation traditionnelle (il s'agit des travaux de Norman McLaren et de Mirai Mizue) on va proposer qu'à la base de toutes ces productions il y a un même dispositif et un même type de pratique (dans le sens sémiotique du terme), un « faire technique » sur lequel on projettera des contenus divers (narratifs, dynamiques, musicaux, énergétiques, etc.). Ainsi, j'espère ouvrir la porte (ou au moins, une fenêtre) au champ de l'animation (dans toutes ses formes) pour le projet sémiotique qui aujourd'hui se trouve lui-même dans une période d'expansion théorique, en découvrant de nouveaux plans d'immanence, où le dernier horizon est la vie même, et ses formes de sens toujours dynamiques.

# Chapitre 1. Le son au cinéma : des études filmiques à la narratologie audiovisuelle

J'ai choisi de commencer par un parcours à travers des théories ou réflexions de différents auteurs sur la question du son au cinéma, parce qu'on trouve toujours dans les traités, recueils et ouvrages spécialisés sur le sujet, la même plainte : qu'il n'y a pas suffisamment d'études sur le son au cinéma. Cependant, même si en comparant les réflexions sur d'autres sujets relatifs à ce moyen d'expression (comme l'image, la narrativité, ou le montage), on peut dire que, effectivement, le son occupe une place périphérique, il y a pourtant eu différents théoriciens (ou écrivains, ou même des cinéastes) qui ont consacré des pages de grande valeur sur cette insaisissable question. En les découvrant, on peut profiter de ces apports, et aussi on peut identifier les vides, les incomplétudes.

Avant de commencer, un avertissement : bien que dans ces pages on va questionner la pertinence d'une séparation radicale de la musique, des paroles et des bruits pour parler de la bande-son, dans cette enquête j'ai mis l'accent sur les derniers, c'est-à-dire, les sons comme tels, avant de (ou indépendamment) qu'ils deviennent paroles ou musiques. Bien sûr, cette réduction méthodologique est devenue intermittente, en raison du fait que, dans de nombreux cas, les réflexions sur le son ne se développent qu'à partir d'analyses ou commentaires sur les deux composants les plus étudiés : les dialogues et les pistes musicales.

Dans les études filmiques, la question du son est apparue après l'émergence du film sonore, datée vers 1927<sup>3</sup>. Néanmoins, des cinéastes eux-mêmes ou les premiers critiques et théoriciens se sont occupés du nouveau phénomène très tôt. Dans le recueil *Film Sound: Theory and Practice*, Elisabeth Weis et John Belton consacrent la deuxième partie à ce qu'ils appellent la *Classical Sound Theory*, où l'on trouve un échantillon des textes de cette première étape (WEIS & BELTON, 1985) : le manifeste d'Eisenstein, Pudovkin et Alexandrov,

---

<sup>3</sup> Bien sur, la sonorité au cinéma ne se limite au son « synchronisé ». Les historiens du cinéma ont exploré largement le sujet des sons à l'ère du muet. Voir la note 135. Et différentes expériences de son synchronisé existaient aussi avant de 1927. Voir notes 155 à 158. En outre, Jean Mitry souligne le fait que le cinéma sonore existe même *avant* le cinéma tout court, car en 1889 W.K.L. Dickson, collaborateur d'Edison, avait projeté dans le kinetograph le premier film avec musique, à travers un phonographe connecté au dispositif des images (MITRY, 1963).

*L'avenir du film sonore*, et le texte de Pudovkin sur l'asynchronisme (1928) ; un texte de René Clair de 1929, *L'art du son* ; le dernier chapitre du livre *Film as Art* de Rudolph Arnheim : *A New Laocoon: Artistic Composites and the Talking Film* ; un texte d'Alberto Cavalcanti, *le son dans les films* ; un chapitre consacré au son par Bela Balazs ; le chapitre 7 de *Théorie du film* de Siegfried Krakauer, intitulé *dialogue et son*<sup>4</sup>; et l'article de Jean Epstein, *Le Ralenti du son* (1948).

Dans ces textes, malgré leur diversité, on trouve une perplexité similaire face au nouveau moyen d'expression (le son et, notamment, la voix et la parole), et les possibilités et difficultés que ses combinaisons avec l'image pouvaient apporter au cinéma : l'idée de contrepoint et d'asynchronisme revendiquée par les Russes, les possibilités expressives de la manipulation du son chez Epstein, la reconnaissance des sons en général, au-delà des paroles chez Clair, les intuitions de Balazs sur le pouvoir expressif des bruits ou la systématisation rigoureuse et anticipatoire de Krakauer<sup>5</sup>...

D'après certains historiens des théories du cinéma comme Francesco Casetti (CASETTI, 2005) ou Robert Stam (STAM, 2000), il y aurait différents paradigmes ou moments dans les réflexions sur ce médium. Casetti identifie trois paradigmes, le premier étant celui qu'il appelle les théories « ontologiques », qui conçoivent le cinéma en relation avec la réalité, c'est-à-dire dans sa dimension référentielle, surtout par rapport à l'image et à l'enregistrement photographique de la réalité. Mais les auteurs du recueil de Weis y Belton qu'ont vient de citer laissent voir qu'on trouve déjà, grâce aux questions ouvertes pour le cinéma sonore, d'autres interrogations au-delà de la référencialité.

Pourtant c'est dans la seconde étape identifiée par Casetti qu'on trouve les propositions les plus systématiques. Mais avant l'émergence de ce groupe, appelé par Casetti les théories *méthodologiques* (dans lesquelles l'auteur comprend la filmologie et la sémiologie du cinéma), il y avait des antécédents chez quelques auteurs qui ont travaillé sur la conception du cinéma comme langage, et qui ont proposé aussi des réflexions sur la dimension sonore. Parmi ces théoriciens les plus connus retenont Jean Mitry et Marcel Martin<sup>6</sup>.

---

<sup>4</sup> On reviendra après sur les idées de Krakauer par rapport à la narratologie du cinéma.

<sup>5</sup> On reviendra sur le travail de Krakauer dans le paragraphe suivant.

<sup>6</sup> Francesco Casetti mentionne quelques chercheurs italiens qui, entre les années quarante et soixante ont travaillé aussi sur les différents « milieux expressifs » que le cinéma utilise, mais il ne parle

## Du rythme au son : Mitry et Martin

Dans le cas de Mitry, dans son chef-d'œuvre *Esthétique et psychologie du cinéma*, on trouve la question du rythme dans la troisième partie du premier volume (MITRY, 1963a). Avec une grande rigueur, Mitry décrit ce phénomène dans la musique et la parole, et après se concentre dans le rythme cinématographique et ses relations avec le montage. D'après Mitry il y a deux tendances : premièrement, les cinéastes et théoriciens ont trouvé dans la musique le modèle pour libérer le cinéma de sa dépendance à l'égard de l'anecdote, d'une théâtralité fautive, en découvrant ainsi une « musique visuelle » à travers le montage. Deuxièmement, les premiers cinéastes abstraits ont privilégié l'image et la construction d'un rythme purement visuel, indépendant des structures musicales. Néanmoins, à son avis les deux ont échoué dans la tentative de trouver un nouveau rythme entre les images et la musique.

Dans le deuxième volume de son œuvre, Mitry reprend le sujet du rythme et d'autres problèmes liés au son, dans le chapitre *la parole et le son* (MITRY, 1963b). L'auteur propose une série de huit étapes par lesquelles le film a conquis sa continuité (ou plutôt, sa propre « durée » narrative) à travers l'histoire, et il place l'incorporation de sons et dialogues dans la septième étape, la naissance d'un véritable art de l'audiovisuel. Pour Mitry, le dialogue relie le cinéma à la littérature, alors il se concentre plutôt sur cet aspect de la bande sonore. Cependant, Mitry trouve que certains auteurs soulignent le rôle des effets sonores et de leur décalage avec l'image : les premiers cinéastes russes, René Clair, André Levinson, parmi d'autres, qui croient qu'il est possible d'explorer des relations multiples entre la bande sonore et la bande visuelle. De même (en périphérie dans son œuvre) Mitry souligne la richesse des possibilités du décalage sonore dans le genre fantastique, la comédie et *les dessins animés*. Précisément, Mitry revient sur ce sujet dans le sous-chapitre intitulé *Structures audiovisuelles*.

*Dans certaines Silly symphonies, néanmoins, l'arabesque visuelle inscrite autour du thème musical des variations de formes et de mouvements, une manière de transposition plastique qui n'est pas toujours sans valeur. Ces courts métrages, le résultat d'une structure rigoureuse et d'une imagination très poétique ont été ballets burlesques qui restent parmi les œuvres cinématographiques les plus valides. Mais si les dessins animés ne sont pas plus que le mouvement des lignes et des couleurs, pour peu que Mickey Mouse soit dessiné à partir*

---

d'aucune étude consacrée au son spécifiquement. Celui qui se rapproche le plus de cette question est Alberto Menarini (MENARINI, 1955), sur les dialogues et la parole au cinéma.

*des graphismes abstraits, ces mouvements peuvent devenir des mouvements de « formes pures » (Mitry, Esthétique et psychologie du cinéma. 2. Les formes, 1963, pp. 166 – 167).*

D'après Mitry, dans ces premiers courts métrages animés, Walt Disney a réussi là où les cinéastes de l'avant-garde ont échoué, en essayant de construire un rythme visuel propre. L'auteur affirme que dans presque tous les cas, c'est le rythme musical qui impose aux images leur métrique. À ce sujet, Mitry souligne le triomphe d'un autre animateur : Alexandre Alexeïeff et son film *Une nuit sur le Monte Chauve* (1933), « un chef-d'œuvre du genre en même temps qu'un chef-d'œuvre tout court ». Mitry souligne la richesse de la technique de l'écran d'épingles et sa capacité pour produire des images « sans lignes », mais avec « un perpétuel jeu d'ombres et de lumières, des formes imprécises, fantomatiques hallucinantes, que l'œuvre de Moussorgsky semblait faire surgir des enfers en les animant de son souffle prodigieux ». Et sa description continue en exaltant les nouveaux rapports entre image et musique :

*Plus de gratuité non plus, mais une émotion plastique prolongeant l'émotion musicale, l'une et l'autre moulée dans un même rythme et constituant son développement même ; une œuvre géniale s'il en fut et qui surpasse les remarquables graphismes de McLaren et les savantes géométries de Fischinger. Et le film, cette fois, peut être projeté sans la musique. Il y perd énormément : sa résonance intérieure, sa temporalité oppressante, tout ce qui lui vient d'elle, cela va sans dire, mais la vision demeure dont les fantômes trouvent, en leur mouvement et leur forme même, une justification suffisante (Mitry, 1963, pp. 353).*

L'intérêt constant pour formaliser les composants de l'expression cinématographique rapproche Mitry du projet sémiotique. Pourtant, cet auteur défiait toujours l'idée que le cinéma était un langage<sup>7</sup>. Au contraire, Marcel Martin est un auteur « pré-semiôtique » qui s'est intégré sans problèmes dans la tradition de la sémiologie filmique<sup>8</sup>. En ce qui concerne le son, il lui consacre le chapitre 7 de son livre *Le langage cinématographique* aux phénomènes sonores (MARTIN, 1955). Une vertu de ce travail est qu'il assume le bruit spécifiquement en tant que producteur de sens. Pour commencer, Martin se souvient, comme Mitry, des cinéastes russes et de leur volonté de créer du sens à partir du déphasage entre le

---

<sup>7</sup> De fait Mitry a consacré tout un ouvrage à contester la pertinence de l'approche sémiologique, et un de ses arguments c'est que la sémiologie a omis la question du rythme (MITRY, 1987, p. 215).

<sup>8</sup> En effet, Marcel Martin se considère comme un sémioticien, et dans la révision de son travail qu'il a fait plus récemment, il a inclus une annexe de terminologie sémiotique.

son et l'image. De même, Martin souligne comment le film muet remplaçait les sons avec des stratégies de montage, et comment l'incorporation du son a transformé les formes narratives. Le son, nous dit Martin, peut être utilisé comme contrepoint ou contraste de l'image, et de manière réaliste ou non réaliste, en produisant quatre modes possibles d'organisation des relations image-son. Par ailleurs, le son peut provenir d'une source interne à l'écran ou d'une source hors du cadre.

Marcel Martin résume les contributions du son au cinéma :

- Un accroissement de l'impression de réalité, de l'authenticité de l'image.
- La continuité sonore, qui rétablit la continuité des images.
- La suppression des intertitres. Libération des images de la fonction explicative.
- Le silence devient une valeur positive, un marqueur dramatique.
- La possibilité d'inclure des ellipses sonores.
- Le décalage entre le son et les images comme mécanisme expressif.
- Le nouveau rôle de la musique comme matériau expressif, quand il n'y a pas une justification diégétique.

Ensuite, Martin se concentre dans le rôle des bruits dans le cinéma et propose une classification élémentaire :

- Les *bruits naturels*, c'est-à-dire, tous les phénomènes sonores qu'on peut percevoir dans la nature : le vent, le tonnerre, la pluie, les vagues, les cris des animaux, etc.
- Les *bruits humains*, parmi lesquels on peut différencier :
  - les *bruits mécaniques* (des machines, voitures, locomotives, avions, bruits des rues, ports, stations...)
  - Les mots-bruits, c'est-à-dire, le fond sonore humain, les murmures, les langues étrangères.
  - La musique-bruit, comme la musique des musiciens ou la musique qu'émettent les radios, etc.

Tous ces bruits peuvent être utilisés d'une façon réaliste, quand le son correspond à la source sonore que l'image montre ou qui a été montrée dans quelque moment du film. Cet usage réaliste des bruits se manifeste, d'après Martin, sous la forme de la synchronisation

avec des images. Mais les bruits peuvent être utilisés aussi d'une façon « irréaliste », à travers le déphasage entre les sons et les images, comme l'ont proposé les cinéastes russes. De cette façon, les bruits peuvent produire métaphores et symboles. Ainsi, la proposition de Martin inaugure une rhétorique du son dans le cinéma<sup>9</sup>.

Tant Mitry que Martin ont prévu ce qui, d'après Casetti, sera le deuxième paradigme des théories du cinéma : les théories dites « méthodologiques », parmi lesquelles figure la sémiotique. L'antécédent direct de cette approche c'est la filmologie, avec une forte influence de la psychologie, mais aussi d'autres sciences humaines<sup>10</sup>. Dans plusieurs œuvres de cette tendance, des auteurs comme Gian Piero Brunetta ont consacré des œuvres entières au son dans le cinéma, bien que tardives (BRUNETTA, 1970)<sup>11</sup>.

Mais le chapitre de l'histoire des théories le plus intéressant pour cette recherche est celui qui correspond à la sémiologie du cinéma, dans la figure de son représentant principal : Christian Metz. D'après Casetti la période correspondant à la sémiologie du cinéma se caractérise par le développement des méthodologies de recherche rigoureuses, appuyées par des institutions universitaires et des centres de recherche. Le cas de Metz est emblématique à

---

<sup>9</sup> Marcel Martin aussi propose une classification pour les usages de la musique dans le cinéma, en distinguant la fonction rythmique, la fonction dramatique et la fonction lyrique. De même, l'auteur mentionne les conceptions différentes de la musique cinématographique : une conception « synthétique » (chez Pudovkin) selon laquelle la musique doit centrer l'attention du spectateur sur une situation comme totalité, à travers le décalage. Et une conception « analytique » (chez Eisenstein) à travers le concept de « contrepoint » audiovisuel, la concordance entre les effets visuels et les motifs musicaux. Dans une révision et réédition de cette œuvre, Martin a inclus un paragraphe sur le concept de « point d'écoute » (proposé par Michel Chion), de la même façon qu'on parle du « point de vue ». Ces deux points coïncident généralement, dans les personnages et les spectateurs. Mais quelquefois, le spectateur peut écouter le son à travers le point d'écoute d'un personnage, de façon subjective ou objective. Ainsi, Martin déploie un ensemble d'oppositions dans une volonté classificatrice aussi proche de la tradition rhétorique des figures.

<sup>10</sup> Voir à ce sujet (JULLIER, 2009).

<sup>11</sup> Bien qu'il soit difficile de trouver des réflexions spécifiques sur le son dans la filmologie, on trouve toujours des métaphores relatives aux phénomènes sonores ou musicaux pour parler des autres dimensions du cinéma. Par exemple, chez Etienne Souriau, on l'« écoute » parler du rapport « d'harmonie » entre les sensations et les perceptions visuelles (SOURIAU, 1947, p. 52).



cet égard. Et bien que la question du son au cinéma ait été toujours périphérique, le sujet n'a pas été complètement indifférent pour le sémiolinguiste.

### **Le son chez Christian Metz**

Dans son célèbre essai, *Le cinéma : langue ou langage ?* Christian Metz évoque la question de l'avènement du cinéma parlant, bien qu'il se concentre sur tout dans la dimension de la parole (METZ, 1964). Metz questionne la peur et la méfiance que le parlant a suscité chez les cinéastes et théoriciens, avec la seule exception de Marcel Pagnol. La citation suivante laisse voir le diagnostic de Metz sur cette époque-là :

*...l'apparition de la parole dans le film n'a pas modifié substantiellement les positions théoriques en présence. On sait que maint amant de la pureté cinématographique se fit un peu prier avant d'admettre la nouvelle venue dans l'univers filmique. On assura çà et là qu'on ne l'utiliserait pas, ou le moins possible, et en tout cas jamais en fonction réaliste, que du reste la mode en passerait bien vite... Il y eut aussi une véritable opération — diversion et retardement à la fois — qui consista à jouer le « sonore » contre le « parlant » : on acceptait les bruits réels, la musique, mais non la parole qui, seule de tous les sons du monde au nombre desquels elle figure pourtant, demeurait frappée — en théorie — d'un mystérieux et spécifique interdit : ce fut « le film parlant sans paroles, le film réticent, le film orné de bruits de porte et de tintements de cuiller, le film gémissant, criant, riant, soupirant, sanglotant, mais jamais parlant » qu'a épinglé avec tant de verve le dramaturge du Midi1. Mais tout cela n'était qu'épisode. Ce vaste débat sur l'admission statutaire de la parole dans le film fut largement platonique : on ne pouvait que reconnaître son inscription dans la théorie du film. Les films eux-mêmes furent parlants, ils le furent très vite, ils le furent presque tous, ils le restèrent (METZ, 1964, p. 67).*

Metz recueille ce refus dans son célèbre « paradoxe du cinéma parlant », décrit comme suit : « Quand le cinéma était muet, on lui reprochait de trop parler. Quand il se mit à parler, on déclara que pour l'essentiel il restait muet et devait le rester. Quoi d'étonnant, dès lors, si l'on affirmait que l'avènement du parlant n'avait rien changé ? » (METZ, 1964, p. 69). À partir de cette idée, Metz continue en rapprochant le cinéma de la langue et du langage, en oubliant la sonorité dans un sens plus vaste.

Pourtant, dès les premiers travaux de Metz, recueillis dans les *Essais sur la signification au cinéma*, cet auteur a parlé du cinéma comme un langage composé par de multiples codes, constitués par des matières de l'expression différentes. Dans ces premières

œuvres il en mentionné quatre, mais dans son œuvre la plus vaste, *Langage et cinéma*, Metz en ajoute une de plus :

*...le discours cinématographique inscrit ses configurations signifiantes dans des supports sensoriels de cinq ordres : l'image, le son musical, le son phonétique des « paroles », le bruit, le tracé graphique des mentions écrites. L'analyse structurale, c'est vrai, s'intéresse à la figure, non à son support matériel ; mais celle-là doit à celui-ci certain de ses caractères, et les deux choses ne sont séparables que jusqu'à un certain point. (...) On ne saurait, en tout cas, définir le film comme fait de langage si l'on refuse de prendre en compte qu'il « joue » sur cinq matières signifiantes, et sur ces cinq-là : dans cette mesure, la sémiologie du film se mêle inextricablement de considérations « psychologiques » (mécanismes perceptifs, caractères propres de l'image, etc.), réutilisées dans une autre perspective. (METZ, Langage et cinéma, 1971, p. 10)*

Metz reconnaît donc les sons comme des éléments cinématographiques, aussi fondamentaux que les images, à partir des mécanismes perceptifs, mais il ne développe pas beaucoup l'idée plus loin dans cette œuvre. Cependant, dans un chapitre ultérieur, en révisant les distinctions hjelmsléviennes entre matière, substance et forme, Metz évoque la question du rythme, en termes similaires à ceux de Mitry :

*Un schème rythmique – même si l'on peut, et même si l'on doit l'analyser comme une pure forme relationnelle, comme un rapport abstrait de grandeurs mesurables – est un phénomène intrinsèquement temporel ; si on l'oubliait, on saisirait encore un schème, mais non plus un schème rythmique. Aussi peut-on dire qu'un tel schème est « propre », sinon à la musique seule, du moins à un ensemble de langages dont la matière de l'expression présente le caractère de s'inscrire dans le temps (et par conséquent, entre autres, à la musique et au cinéma, ce qui rend possibles les remarques d'Eisenstein tout en interdisant de les généraliser). (METZ, 1971, p. 166)*

Finalement, dans le dernier chapitre de *Langage et cinéma*, consacré aux rapports entre cinéma et écriture (METZ, 1971, pp. 209 – 217), Metz revient sur le son, quand il établit un parallèle entre cinéma et idéographie. L'auteur souligne ici certaines caractéristiques du son :

*Les technologies cinématographiques autorisent des reproductions sonores d'une plus haute fidélité que les reproductions visuelles, et donc d'un degré moindre de schématisation ; le fonctionnement de la perception filmique et celui de la perception réelle se ressemblent davantage pour le son que pour l'image ; à l'écoute « naïve », il y a peu de différence*

*phénoménale entre le son enregistré et le son directement entendu, si du moins les techniciens s'en sont donné la peine. Les modes de schématisation interviennent aussi. Dans l'image, certains éléments son ouvertement codés : le rectangle écranique, l'absence de ce que les psychologues appellent la « troisième dimension réelle » (= binoculaire), et dont la régulière supplétion par l'impression dite de profondeur (= perspective monoculaire+mouvement) n'échappe pas à celui même qui supplée sans effort, etc. ; les faits de genre, qui ont été étudiés de près par la filmologie et par certains esthéticiens du cinéma, vont l'encontre de l'impression de réalité ; or il n'existe pas, en matière de son filmique, de distorsions de même importance (cela tient en partie à ce que l'ancrage spatial d'un son est plus faible et plus vague que celui d'une donnée visuelle). (Metz, 1971, pp. 209 – 210)*

Ainsi, Metz revient à la question de l'expérience audiovisuelle et à sa relation avec les marques textuelles, en anticipant un tournant pragmatique qui permet de lier les « énoncés » filmiques avec les expériences des spectateurs, comme l'ont fait Roger Odin, dans sa proposition sémio-pragmatique<sup>12</sup>, et les narratologues du cinéma, que nous allons revisiter dans le chapitre 2.

Mais on ne peut pas laisser Metz sans mentionner son essai *Le perçu et le nommé*, consacré spécifiquement au sujet du son (METZ, 1977, pp. 130-161)<sup>13</sup>. Au début, l'auteur se concentre sur les relations entre la langue et la perception visuelle, et surtout dans la dimension lexicale qui s'occupe de nommer les objets. Metz propose à cet égard deux conclusions :

---

<sup>12</sup> Comment dit Warren Buckland dans son œuvre *The Cognitive Semiotics of Film*: « le terme «mode» d'Odin englobe les divisions que Metz inclut dans la catégorie des *textes filmiques* –classes et groupes de films. Cependant, alors que les catégories de Metz restent résolument taxinomiques, les modes d'Odin sont cognitifs. Ils se réfèrent à une partie importante de la compétence discursive de chaque spectateur. L'espace de communication est créé lorsque l'émetteur et le récepteur utilisent la même compétence –le même mode– dans la production et la lecture d'un film » (BUCKLAND, 2004, p. 88). (*“Odin’s term ‘mode’ encompasses Metz’s divisions within the category of filmic texts – classes and groups of films. However, whereas Metz’s categories remain resolutely taxonomic, Odin’s modes are cognitive – they refer to a major part of each spectator’s discursive competence. The space of communication is created when sender and receiver employ the same competence – the same mode – in the making and reading of a film”*).

<sup>13</sup> Une partie de ce texte de Metz a été inclus dans le dossier de *Yale French Studies* édité par Rick Altman, sous le titre en anglais d'*Aural Objects* (METZ, 1980).

Il y a une équivalence entre le « code linguistique » et le « code perceptif objectif » au niveau des signifiés (mais pas au niveau des signifiants). Cette équivalence sémantique est entre un « objet reconnu » et un « sémème », comme décrit par Greimas dans sa *Sémantique Structurale*.

La langue est un « métacode », qui rend compte du « code objectif » de façon que la relation entre les signifiés des codes soit motivée, mais le rapport entre les signifiants linguistiques et les signifiés perceptifs est arbitraire. D'après Metz, la sémantique lexicale est plus vaste que la sémantique perceptive : « la langue peut dire, même si c'est parfois avec approximation, ce que disent tous les autres codes, alors que l'inverse n'est pas vrai dans nulle mesure » (METZ, 1977, p. 148).

Ensuite, Metz s'occupe des bruits et des « objets sonores », en proposant des conclusions précises :

Premièrement, pour Metz, la culture occidentale est essentialiste, dans la mesure où elle fait la distinction entre qualités primaires (essentielles) et secondaires (non essentielles). En général, les sons sont considérés comme des qualités secondes, c'est-à-dire des attributs des objets, des «essences». C'est pourquoi on doit ajouter l'adjectif « sonore » à ce type d'objets, tandis que les objets visuels (un drapeau, par exemple), sont appelés tout simplement objets. Ceci parce que selon cet essentialisme les qualités visuelles et tactiles sont souvent considérées comme «essentielles» et donc suffisantes pour identifier un objet, tandis que les qualités du son ou même de l'odeur sont considérées comme secondaires, juste des attributs des objets qui peuvent être reconnus sans eux.

Cependant, il y a des expériences sonores qui sont désignées par des adjectifs dans la langue. Par exemple, le "sifflet" est un nom, et il est donc de droit un objet. Cependant, l'«essentialisme» dont parle Metz a fait que de tels phénomènes soient considérés comme des «quasi-objets».

D'où une erreur commune au cinéma, qui vient du vocabulaire technique : le « sound-off » (y compris les voix et la musique) ne l'est pas vraiment :

*On oublie que le son lui-même n'est jamais « off » : ou il est audible, ou bien il n'existe pas ; quand il existe, il ne saurait pas être situé à l'intérieur du rectangle ou en dehors, puisque le propre des sons est de diffuser plus ou moins dans tout l'espace environnant : le son est à la*

fois « dans » l'écran, devant, derrière, autour, dans toute la salle du cinéma. (METZ, 1977, pp. 157-158)

Le terme « sound » signifie ici en réalité "sources sonores" qu'on voit à l'écran ou non. Encore une fois, on essaie de saisir un phénomène flou, immatériel, "quasiobjectual" à partir de l'ancrage dans une essence visuelle, pour laquelle le son ne serait qu'un attribut. La confusion vient du « langage des techniciens et des studios », qui « sans s'en apercevoir, applique au son une conceptualisation qui n'a de sens que pour l'image : on prétend parler du son, on pense en fait à l'image visuelle de la source sonore » (Ibid.).

Christian Metz conclut son texte avec une idée très intéressante, et qu'on va retrouver dans les développements contemporains de la sémiotique : « Il me semble pour ma part que l'entreprise sémiologique tout entière, à travers son ancrage initial dans le souci du signifiant perceptible, de ses écarts perceptibles, etc., s'inscrit d'une certaine façon dans le prolongement de l'inspiration phénoménologique » (p. 159). Aujourd'hui cette idée trouve des résonances dans la sémiotique post-greimassienne où dans la sémiotique cognitive, deux courants éloignés par autres aspects. Pour finir partiellement ces brèves lignes sur le fondateur de la sémiologie du cinéma, on va voir que dans le dernier Metz la question des paroles, des sons et de la musique sera intégrée à son idée de l'énonciation impersonnelle, comme partie des stratégies métadiscursives du cinéma (METZ, 1991).

## **Le son dans la narratologie du cinéma**

La question du son dans le cinéma a été largement discutée. Depuis le début des réflexions dans la période de transition du muet au parlant<sup>14</sup>, jusqu'aux conceptualisations du son multipistes et la numérisation du son d'aujourd'hui, le rôle des bruits contribue au débat sur le sens au cinéma, plus ou moins périphérique selon les auteurs et les tendances. Dans une première étape, la relation entre le son et l'image a été structurée selon sa coïncidence (ou non-coïncidence). Mais il n'était pas toujours évident que les différents auteurs parlent des différences dans le plan d'expression ou dans le plan du contenu. Pierre Schaeffer, un des premiers théoriciens rigoureux du son, a parlé de quatre types d'effets dans les relations images/sons : l'effet *masque* (l'image masque la musique) ; l'effet d'*opposition* (quand la

---

<sup>14</sup> Mais aussi avant, après les accompagnements de la musique et des narrateurs, et même des effets de bruitage ont été toujours courants.

musique dit quelque chose de différent que l'image) ; l'effet *synchronisme* (lorsque le rythme musical s'accorde très précisément avec celui de l'image) ; et l'effet de *syntonie* (quand des rythmes différents –visuels et sonores– réagissent l'un sur l'autre, en générant une résultante plus forte que la simple addition des deux impressions originaires). Les deux premiers semblent parler des relations sémantiques, et les deux secondes des relations expressives, mais le premier peut aussi parler des relations expressives et le quatrième, des relations sémantiques.

En tout cas, les différentes propositions des auteurs peuvent se décliner dans une de ces options, ou leur combinaison. Ainsi, Sigfrid Kracauer, qui, de l'avis de Roger Odin inaugure les conceptualisations les plus systématiques de la théorie du cinéma, fait la distinction entre six axes :

**Axe 1 : code des relations de prédominance.** Il s'agit de savoir si dans un film les messages passent essentiellement par l'image ou par la bande-son.

**Axe 2 : code du synchronisme : synchronisme vs asynchronisme.** Kracauer propose de réserver le terme de « synchronisme » aux sons et aux images qui, coïncidant sur l'écran, sont aussi synchrones dans la vie réelle, c'est-à-dire aux sons « qui peuvent être enregistrés par une caméra à prise de son synchrone ». Dans tous les autres cas, on parlera d'asynchronisme.

**Axe 3 : code des relations sémantiques : parallélisme vs contrepoint.** Sons (paroles) et images peuvent exprimer des significations parallèles, contradictoires, ou indépendantes les unes des autres.

**Axe 4 : code de la relation des sons au monde du film: son réel vs son commentatif (*actual sound vs commentative sound*).** Les premiers sont des sons qui appartiennent au monde représenté sur l'écran; les seconds sont des sons dont la source n'appartient pas au monde représenté sur l'écran. (Ce qu'on appelle plus tard la distinction diégétique/extra-diégétique).

**Axe 5 : code de l'identification des sources sonores : sons reconnaissables vs non reconnaissables,** (Selon qu'ils permettent ou non l'identification de leur source).

**Axe 6 : code du réalisme sonore : sons réalistes vs non réalistes.** Le réalisateur est libre de détourner les sons de leur ancrage naturel et d'associer à une image des sons tout à fait

différents de ceux que l'on attendrait normalement; Kracauer indique que cette opération confère souvent au son une valeur symbolique.

Selon Roger Odin, après Kracauer il y a eu peu de nouveau, et toute la tradition de la sémiologie et de la narratologie du cinéma n'a fait que développer et détailler ces concepts. Et c'est parce que ces axes permettent de regarder la relation entre expression et contenu qu'ils ont été assimilés par les approches sémiotiques. Principalement, les axes deux, trois et quatre ont été les plus développés, c'est-à-dire, la question du synchronisme, la question des rapports sémantiques entre sons et images et la distinction entre les éléments qui appartiennent au monde diégétique et ceux qui lui sont extérieurs (ce que, dans sa terminologie pré-narratologique, Kracauer appelle « sons réels vs sons commentatifs ». La seule distinction qui soit apparue après, de l'avis d'Odin, est le couple « son in »/« son off », qui a une origine plus technique que théorique, parce que ce qui est « in » ou « off » c'est la source des sons dans la représentation à l'écran. Le problème avec cette distinction est, précisément, que la notion de « champ » est purement visuelle, par rapport au moment de la prise de vues, ce qui a produit beaucoup de malentendus (comme l'ont signalé plusieurs auteurs, il n'existe pas un « hors de cadre » sonore).

Après Kracauer, il y a eu plusieurs propositions de taxonomie ou de systématisation des rapports son/image. Parmi les plus connus : D. Percheron (1973), A. Gardies (1980) et D. Château et F. Jost (1979) et D. Château (1986)<sup>15</sup>.

Roger Odin, lui-même, n'échappe pas cette tradition, bien qu'il essaie de proposer une nouvelle approche, « en termes d'inférence ». Odin va développer sa théorie autour de trois questions : le synchronisme, qu'il développe à partir du concept de « rythme » ; la diégétisation des sons, pour laquelle l'auteur propose l'idée d'une « compétence encyclopédique sonore », une « compétence encyclopédique visuelle », qui constituent le *code sémantique énonciatif*, ainsi qu'un « code de désignation des sources sonores dans le monde » ; et la question de la construction des sources sonores, qui cherche à résoudre le problème de la distinction son *in* / son *off*, ainsi que la question générale de la construction des sources sonores. À cette fin Odin propose une nouvelle typologie :

---

<sup>15</sup> Pour une recension de ces autres propositions, voir ODIN (1991), chapitre 10, et JULLIER (1995), chapitres 3 et 4.

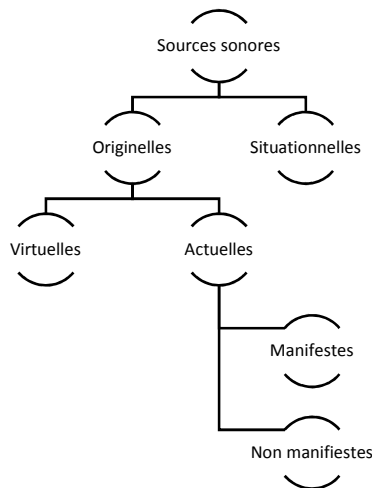


Figure 2 : Typologie des sources sonores

Source : [ODIN, 1990, p. 254 (adapté)]

Les sources sonores originales sont celles qu'on peut identifier grâce à la compétence encyclopédique sonore, tandis que les sources situationnelles sont des objets susceptibles de devenir des sources sonores dans quelques situations. Les sources sonores actuelles seront manifestes lorsqu'elles manifestent par un mouvement visible dans l'image leur statut de source en train d'émettre.

La particularité de l'approche d'Odin par rapport à la tradition sémiologique et narratologique est que ces précisions tournent autour de « l'écoute filmique », dans le besoin de prendre la position du spectateur : « Ce qu'il faut comprendre, ce sont les *inférences* que le spectateur doit effectuer pour comprendre les relations images-sons ».

La principale raison pour adopter cette approche est la particularité des conditions de l'écoute filmique : la première caractéristique est que le spectateur est motivé : « son oreille est prête à pratiquer une *écoute sélective*, c'est-à-dire à effectuer les inférences nécessaires pour distinguer les sons *filmiques* (les sons qui appartiennent au discours signifiant du film) des sons *a-filmiques* ».

Une deuxième question est rapportée aux caractéristiques des sons filmiques : ils sont « acousmatiques et relayés ». La question est, donc, comment le spectateur met en relation ces émissions sonores, désancrées de leurs sources originales (et doublement désancrées puisque le spectateur ne voit pas les « vraies » sources sonores : les hautparleurs), avec le monde



narratif du film. C'est pour affronter ces trois questions qu'Odin propose ces trois types de « codes », maîtrisés par le spectateur, pour mettre en rapport des images et des sons.

Dans son œuvre *Les sons au cinéma et à la télévision*, Laurent Jullier apporte quelques compléments à cette proposition, et ajoute d'autres outils descriptifs intéressants. Au lieu de l'opposition entre sources sonores actuelles manifestes/non manifestes, il propose de parler de l'*inélucltabilité* de la manifestation sonore accompagnant la manifestation visible :

*Il est fort possible, en effet, de donner toutes preuves visuelles du caractère manifeste de certaines sources sonores sans que celles-ci deviennent le moins du monde actives (faire semblant de parler, de jouer du saxophone ou d'allumer un récepteur de radio). Pour d'autres sources, en revanche, cela ne se conçoit pas : impossible de voir abattre un marteau, tomber un verre ou s'agiter les branches un jour de grand vent sans savoir qu'inélucltablement, qu'on l'entende ou non, un son sera produit qui accompagnera ces manifestations visuelles. (JULLIER, 1995, p. 63)*

Cette distinction sert à la détection de ce que Château appelle les *occurrences vides* d'un son filmique (absence d'un son dont la source suggérée émet visiblement). Et pour Jullier il y a aussi des sources actives manifestes pour lesquelles il n'est pas possible d'établir une relation de synchronisme (c'est le cas de la « voix *through* » de S. Daney).

En plus de proposer une combinaison des propositions descriptives qui recueille tout la tradition de la sémiologie du cinéma (mais, principalement, celles de Château et d'Odin), Jullier développe de façon plus forte la distinction diégétique / non diégétique, à partir du concept des « mondes engendrés par l'œuvre ». Cela lui permet d'introduire dans sa description les différentes manifestations du son par rapport à l'activité interprétative du spectateur :

*Les éléments de la bande-son peuvent aller habiter différents mondes : le monde narratif, les hétéro-univers, les fosses, et les mondes de la fabrication et de la projection (ou de la diffusion s'il s'agit d'une émission TV). On ne dira donc pas qu'un son est diégétique ou extra-diégétique, mais qu'il se situe quelque part le long de cet axe qui va du strict monde narratif (des quatre cités, le plus proche de la traditionnelle catégorie du « diégétique ») au monde de la projection, le plus éloigné du « diégétique ». (Ibid., p. 103 – 104)*

Ce qui est commun à ces deux théoriciens est, donc, le choix de la position du spectateur et de son activité d'écoute. Les deux essayent de sortir de l'immanence du « texte

filmique », pour confronter le caractère évènementiel du son et de l'expérience concrète qu'il produit. On pense immédiatement à l'énonciation, une question qui traverse les théories de ces auteurs, mais qui n'est pas assez claire dans aucun d'eux. Mais on peut dire que la question de l'énonciation est aussi présente (et pas plus claire) dans toute la tradition des études autour du son au cinéma. Tant le « synchronisme » que le « contrepoint », la présence ou l'absence des sources dans le cadre, ou l'emplacement « diégétique » des sons supposent des relations entre un « énonciateur », un « énoncé » et un « énonciataire », qui n'ont pas été explicitées. Il est donc nécessaire de confronter ces propositions avec la théorie de l'énonciation, en prenant en compte les particularités qu'une telle théorie doit affronter quand elle est appliquée au cinéma.

### **Michel Chion, ou l'art de la dénomination**

On ne peut pas finir ce parcours sans s'arrêter sur l'œuvre du chercheur qui a travaillé sur le son de façon la plus soutenue et rigoureuse. Dans ses recherches Michel Chion a rendu compte de presque tous les domaines relatifs à la sonorité, dans le cinéma et en dehors du cinéma, en employant tout type d'approches et instruments théoriques, provenant de sources si diverses comme la musicologie, les théories de la perception, la narratologie, la sociologie, l'histoire du cinéma et de la musique, et beaucoup plus, mais avec un sens d'unité et de cohérence, obtenues grâce au point de pivot que lui donne son objet. Chion s'est caractérisé par un sens aigu pour identifier des phénomènes très précis, et par sa capacité de les nommer avec des néologismes efficaces. Ainsi, depuis son ouvrage sur *La voix au cinéma* (CHION, 1982), où il a forgé le terme d'*acousmètre* (ou *être acousmatique*), à *l'audio-divisionnel* de sa dernière œuvre consacrée au son au cinéma (Chion, 2003), Chion offre autour de 100 concepts qui ont pris la forme d'un glossaire<sup>16</sup>, peut être le meilleur format pour son travail, qui a la diversité d'une encyclopédie.

Il est curieux (mais pas complètement erroné) de constater que le travail de Chion a été classé comme une œuvre sémiotique aux États-Unis. Robert Stam et ses collègues (STAM, BURGOYNE, & FLITTERMAN-LEWIS, 1999, pp. 61 - 64), dans leur travail *New*

---

<sup>16</sup> Inclus dans la postface à la dernière édition de *L'audio-vision* (CHION, 2013, pp. 185-229), selon un ordre logique et thématique, et aussi, en format numérique, dans son site internet : <http://michelchion.com/texts>. Dans *Un art sonore : le cinéma* on trouve un autre glossaire, cette fois selon un ordre alphabétique (CHION, 2003, pp. 411-438).

*Vocabularies in Film Semiotics*, le placent dans le chapitre de *Cine-Semiology* sous le titre : *The semiotics of film sound*. Effectivement, beaucoup de ses idées touchent des questions sémiotiques, bien que lui-même ne se considère pas comme un sémioticien.

Dans ma recherche je ferai appel plusieurs fois à ses concepts. Bien sûr, le travail dont procèdent la plupart des idées que je reprends est *L'audio-vision*, où Chion identifie ce type d' « attitude perceptive » que je voudrais spécifier pour le cinéma d'animation. J'y reviendrai bientôt. Maintenant, il faut signaler que cet auteur a travaillé aussi sur le son comme un sujet autonome, en donnant de clés importantes pour une description sémiotique du son. Dans son livre *Le son, traité d'acoulogie* (Chion, *Le son, traité d'acoulogie*, 2010), spécialement, cet auteur a suivi les traces de son maître Pierre Schaeffer en essayant de proposer une théorie complète des objets sonores et de l'expérience d'écoute<sup>17</sup>. Chion prend en compte un large éventail de questions touchant au son : les controverses scientifiques sur la description du son en termes physiques et acoustiques ; la difficulté de définir le son comme un objet, sa possible « réification » ; les paysages sonores, la distribution du son dans l'espace, les scènes sonores ; les relations entre la voix, la musique et le son ; les termes linguistiques et les mots pour évoquer et décrire les sons ; la manipulation technique des sons ; « le couplage audiovisuel » ; la description et la classification des sons (à partir des idées de Schaeffer), les différents types d'écoutes et la délimitation des objets sonores ; la morphologie de l'objet sonore ; et un paragraphe très intéressant sur des questions qui vont au-delà de la théorie de Schaeffer pour décrire le son : la conception dynamique du son d'après François Bayle, inspiré par les idées de Paul Klee sur la forme. Selon Chion, Schaeffer avait immobilisé le son afin de le comprendre, mais sa réalité est dynamique, parce que le son est une énergie en mouvement, comme a dit Bayle. C'est ainsi qu'il est possible d'identifier des « logiques énergétiques » pour décrire des relations actives de transmission, d'échange et de rétroaction entre les sons<sup>18</sup>. Comme son maître Pierre Schaeffer dans son travail de pionnier *Traité des objets musicaux*, Michel Chion offre un ensemble d'idées et de données clés pour la description du son comme un objet sémiotique.

---

<sup>17</sup> Une autre œuvre de Chion consacré à ce sujet est (CHION, 1983), où l'auteur présente de façon didactique le travail de Pierre Schaeffer.

<sup>18</sup> Je reviendrai sur ce sujet dans les conclusions de ce travail.

En ce qui concerne ses travaux sur le cinéma, on peut identifier différentes découvertes et stratégies de description. Ainsi Michel Chion identifie (avec quelques exemples) :

- Une série de phénomènes perceptifs et cognitifs à la base de l'expérience des spectateurs (synchrèse, aimantation spatiale du son par l'image).
- Les différents rapports son/image (audio-vision, contrepoint, point de synchronisation).
- Un ensemble d'effets audiovisuels ou conventions construits par le cinéma tout au long de son histoire (le rendu, la valeur ajouté).
- Les rôles spécifiques des trois phénomènes sonores au cinéma (voix/paroles, musiques, bruits/paysages sonores), mais aussi ces entrecroisements ou ces frontières diffuses (sa définition du cinéma « parlant » comme visuo-logo-visuel).
- Des variations des rôles et d'usages des composants sonores dans différents types de genres, medias, auteurs (visu-audition).
- Des aspects techniques ou technologies qui déterminent aspects esthétiques ou narratifs (super-champ du son multipistes, effet de coulisse).
- Des clés descriptives sur les rôles narratifs et expressifs des sons au cinéma (point d'écoute, les trois frontières, la scène audio-visuelle, etc.).

Ces stratégies de description (et beaucoup d'autres) seront exploitées ici en essayant de trouver quelques applications à l'animation. À ce sujet, on peut commencer par un des concepts pertinents pour la description des phénomènes sonores dans le cinéma animé. Il s'agit du concept du *Rendu*, qui apparaît dans *La toile trouée* (CHION, 1988), mais qui est développé dans *L'audio-vision*. Précisément dans cet ouvrage Chion remarque l'intérêt de ce concept pour décrire ce qui se passe dans le dessin animé, au début à partir du film *Qui veut la peau de Roger Rabbit* (Zemeckis, 1988). En se demandant comment les « artisans du son » de cette production ont abordé la sonorisation du film, Chion décrit :

*...nous avons affaire ici à un bruitage qui tente de donner discrètement une consistance matérielle à un être graphique. Les bruits corporels des toons y demeurent légers, et le seul*

*moment du film comportant un effet sonore destiné à être remarqué comme tel par le public est celui où la pulpeuse Jessica se frotte contre le détective humain joué par Bob Hoskins. Lorsque le crâne bien concret de ce dernier heurte la volumineuse poitrine dessinée, on entend un « chtongg ! » sonnante le creux qui fait rire la salle. Mais il y a bien d'autres moments du film où les toons, et surtout Roger le lapin, font de subtils bruits de frottements et de contacts, évoquant une matière fine, creuse et élastique, analogue à du plastique gonflé d'air. Ainsi, par le son, les bruiteurs de Roger Rabbitt nous apprennent-ils que les toons sont des êtres creux et de faible poids. Si le spectateur ne prête pas spécialement attention à ces sons, cela ne l'empêche pas d'ouïr ceux-ci, et de se laisser influencer par eux dans sa perception des images. Devant l'écran, il croit alors simplement voir ce qu'en fait il entend-voit, selon le phénomène que nous avons qualifié de valeur ajoutée (CHION, 2013, pp. 103-104).*

Après ce commentaire de *Roger Rabbitt*, Chion consacre deux ou trois pages au cinéma d'animation (et particulièrement au *Cartoon*) à partir des résultats d'une enquête engagée par François Delalande, Bernadette Céleste et Élisabeth Dumaurier sur les bruits que font les enfants dans leurs jeux (DELALANDE, et al., 1982/1996). Ces chercheurs ont observé que les « effets sonores » des enfants ne cherchent pas à imiter les « vrais sons » des objets, mais qu'eux imitent *leurs trajets et mouvements*, dans une volonté d'isomorphisme avec le dynamisme des choses. D'après Chion, c'est le même cas dans la conception sonore des *cartoons* :

*Reprenons le célèbre et banal procédé qui consiste à accompagner la montée d'une pente ou d'un escalier par une figure musicale ascendante..., bien que le pas d'une personne qui monte ne gravit pas lui-même l'échelle des hauteurs. Ce qui est imité ici, c'est donc, selon un symbolisme spatial universel des degrés musicaux, le trajet et non le son du trajet. La plupart des déplacements – et le cinéma d'animation en comporte souvent – sont sonorisés de cette façon (CHION, 2013, p. 106).*

Chion mentionne ici le *mickey-mousing* comme un autre exemple de cet isomorphisme entre le mouvement et la musique<sup>19</sup>, mais son dernier exemple est consacré à Tex Avery. D'après Chion, le dynamisme de ses films s'altère si on les voit sans le son :

---

<sup>19</sup> Dans *La musique au cinéma* Chion avait déjà parlé sur les *Silly Sinphonies* à partir des concepts de *synchronèse* et de cinéma synchrono-cinématographique, selon lesquels tout objet et ses mouvements peuvent-être associés avec la musique (CHION, 2003).

*Les figures visuelles muettes se télescopent, s'impriment mal dans la perception, elles courent trop vite. En raison de la relative inertie de l'œil et de sa paresse comparées à l'agilité de l'oreille pour identifier et mémoriser des figures mobiles, le son est l'imprimeur des sensations visuelles rapides, et c'est pourquoi il s'occupe plus des trajets (dont il appuie la perception visuelle) que de sa matière propre (CHION, 2013, p. 107).*

Ainsi trouve-t-on dans le vocabulaire du son forgé par Chion un ensemble de concepts que permettent décrire des aspects de la relation audiovisuelle que prennent les manifestations spécifiques dans le cinéma d'animation, particulièrement, les concepts de *synchrèse* et de *rendu*. Ce dernier est de grand intérêt en ce qui concerne les univers narratifs de l'animation par rapport à ce que Roger Odin appelle la *diégétisation*, tandis que le premier nous servira pour identifier un autre type de combinaison audiovisuelle, dans les animations dites « abstraites ». Dans les deux cas, la structure énonciative des films animés vient au premier plan, dans un phénomène qui sera mieux expliqué d'après une théorie des pratiques sémiotiques.

Pour l'instant on doit souligner que les propositions de Chion ouvrent la perspective de la réflexion sur les sons en introduisant des phénomènes qui étaient exclus de l'approche sémiotique parce qu'ils ne faisaient pas partie de la dimension exclusivement textuelle des films. Dans cette même ligne d'ouverture se situe le travail de Laurent Jullier que nous avons déjà évoqué en ce qui concerne l'emplacement narratif des sons.

### **Laurent Jullier et le tournant cognitif**

Depuis sa thèse de doctorat de 1994, *Le bruit au cinéma. Approche sémio-cognitive des occurrences de bruit dans les films sonores de fiction*, Laurent Jullier a développé une approche cognitive de la question du son au cinéma, qui a comme résultat un ensemble d'œuvres qui ont consolidé cette perspective dans le panorama français. Cette approche lui a permis de s'occuper des dimensions traditionnellement exclues des enquêtes plus proches de la sémiologie du film.

Dans son œuvre *Les sons au cinéma et à la télévision* Jullier présente un parcours général de ces dimensions : une caractérisation des processus cognitifs de l'écoute et le traitement des données sonores ; une description des techniques de l'enregistrement du son, sa transmission et l'impact dans la perception du cinéma ; une problématisation du dispositif audiovisuel, notamment des opérations d'ancrage des sons avec les sources visuelles que les

conventions du cinéma et les processus cognitifs leur attribuent, et du phénomène du synchronisme, en montrant les relations entre les deux concepts ; une description de la scène sonore de la salle de cinéma, et des transformations du son dans cette forme de spatialité. Et une révision de la dimension narrative et des rôles traditionnels et contemporains du son dans sa constitution. En plus, Jullier offre des analyses pratiques des films et de productions télévisuelles pour montrer l'effectivité des outils théoriques développés tout au long de son œuvre. Le grand mérite de cet ouvrage est sa capacité de rendre compte des aspects techniques et psychophysiologiques en montrant leur impact dans les dimensions traditionnellement traitées par les théoriciens du cinéma.

Je voudrais mentionner particulièrement les idées de Jullier, développés dans le chapitre 3, consacré aux *sons dans le dispositif audiovisuel*, sur lesquelles je reviendrai après. Le premier concept qui m'intéresse est celui d'ancrage, que j'ai déjà mentionné à propos de la question des sources sonores.

D'après Jullier, la connexion que le spectateur fait entre des sons et des images est presque inévitable, car elle dépend d'une «illusion», produite par les systèmes perceptifs et cognitifs qui d'après ses propres règles, produisent des impressions non véridiques du monde. Il s'agit de processus *bottom-up* sans intervention de la conscience, automatiques. Il s'agit, d'après Jullier, d'une illusion imposée, dérivée des processus évolutifs d'adaptation. Néanmoins, dans le cas du cinéma dit « classique » l'homogénéité audiovisuelle est *voulue*.

En ce contexte, l'*ancrage*<sup>20</sup> audiovisuel, existe avec indépendance du système technique de connexion entre sons et images. Jullier propose au moins 6 différentes combinaisons techniques, de la musique en direct du cinéma muet, aux possibilités du cinéma multipistes d'aujourd'hui. Et bien qu'il y ait toujours des décalages entre la disposition technique des sons et la façon dont les spectateurs réels les perçoivent, il y a toujours une intégration mentale entre les sources sonores (les haut-parleurs) et les sources lumineuses. Néanmoins, le cinéma peut proposer des ancrages « conditionnés » en différents degrés. Dans du cinéma narratif classique, l'auteur souligne qu'il

*...ne conteste pas le caractère unitaire du dispositif audio-visuel, les sources et espaces réels doivent être oubliés au profit des sources et espaces suggérés, montrés ou non ; le « comment*

---

<sup>20</sup> « Relation de causalité actuelle établie mentalement par l'auditeur entre un son entendu et une source sonore vue » (JULLIER, 1995, p. 185).



*c'est fait » ne doit pas revenir à la surface, et les identifications de sons hors-contexte n'ont pas la moindre importance. (...) Le film narratif classique et le dispositif limitent donc les possibilités de lecture d'un son répliatif donné à un seul mode (JULLIER, 1995, p. 62).*

Cela ne signifie pas que l'on ne puisse pas transgresser cette forme d'ancrage traditionnelle. Effectivement, comme on verra dans le paragraphe suivant, dans l'animation (et, à nouveau, surtout dans le cartoon américain) ces transgressions font partie du « contrat » audiovisuel, ce qui permet qu'on écoute le bruit d'une locomotive quand un animal court.

Quant au synchronisme, Jullier spécifie : « le synchronisme audio-visuel tel qu'on l'entendra ici est un *effet* produit chez le spectateur, un effet soumis aux données de l'écran (*data driven*), et mobilisant les processus ascendants qui pilotent les groupements temporels. Les acousticiens en parlent volontiers comme du *synchronisme subjectif* » (Ibid., pp. 63-64).

Le cinéma peut jouer avec ce synchronisme subjectif avec indépendance de l'ancrage. Pour en décrire les possibilités, Jullier propose revenir sur les notions de Chion, mais en les donnant « définitions plus strictes et plus exclusives que les siennes ». Ainsi on peut identifier :

- Des ancrages opérés sans l'aide du synchronisme.
- Des *synchrèses* (ancrages+synchronisme)
- Des *points de synchronisation* (correspondances de seuils ou rythmes entre séries ou segments en dehors de tout ancrage).

Jullier résume ainsi les rapports entre les possibilités expressives de l'ancrage et du synchronisme par rapport aux processus cognitifs :

*Les opérations mentales qui concernent l'ancrage des sons filmiques oscillent entre deux modèles, le « circuit court », où les automatismes qui interviennent dans la perception du monde peuvent fonctionner (caractère « naturaliste » du dispositif cinématographique, faible importance des conventions, illusion imposée) ; et le « circuit long », où les automatismes qui interviennent dans la perception du monde ne peuvent plus fonctionner (caractère non-naturaliste du dispositif cinématographique, grande importance des conventions) (Ibid. p. 67).*

Jullier fait aussi une description du dispositif de l'ancrage dans une perspective historique (y compris l'histoire des technologies) sur laquelle je reviendrai dans le chapitre 3.



Dans son livre *Analyser un film* (JULLIER, 2012), Jullier revient sur ces questions, et aussi sur d'autres idées développées dans *Cinéma et cognition* (JULLIER, 2002). Il propose une distinction « métaphorique », selon laquelle l'écoute opère comme un dispositif de *surveillance*, tandis que la vue opère comme un dispositif de *vissée*. D'où l'idée que construire la sensation d'un monde englobante demande un point d'écoute vague, tandis que les plans subjectifs demandent l'uniponctualité sonore (ils sont fréquemment accompagnés des bruits intimes du corps).

Dans ce livre Jullier propose quatre types d'écoute :

- L'écoute *esthétique* : se concentre dans la matière sonore elle-même.
- L'écoute *technologique* : porte à considérer les moyens techniques qui portent le son jusqu'à nos oreilles.
- L'écoute *cynégétique* : se met à la chasse d'indices, tant des sources cachées comme des intentions, attitudes, etc. d'un personnage ou d'un environnement.
- L'écoute *structurale* : « vise à repérer les unités discrètes d'un langage (...) la matière sonore n'est intéressante qu'en tant qu'elle désigne un plus petit élément de sens au sein d'un système qui lui est extérieur » (JULLIER, 2012, p. 248).

Jullier nous rappelle que dans le cinéma, comme dans l'expérience quotidienne, la séparation de l'écoute et de la vue est artificielle, parce que notre cerveau est configuré pour intégrer ou reconnaître les sons et les images comme faisant partie d'un même phénomène, qui n'est pas une simple addition son+image, mais une réalité autonome. Il mentionne le *recoupement intermodal*, expression des cognitivistes pour décrire cette tendance d'intégration qui s'explique facilement par son utilité évolutive, tant qu'elle permet de profiter l'omnidirectionnalité de l'écoute pour la reconnaissance des objets et êtres dans le monde. Et le cinéma sonore a su exploiter cette tendance.

Cela ne veut pas dire qu'on ne puisse pas reconnaître des particularités des sons au cinéma, mais il faut « activer » avec une particulière intensité les écoutes esthétiques et technologiques, afin de reconstruire la façon dont un certain son a été fait (surtout dans le cas du genre de la science-fiction).

Finalement, Jullier remarque trois tendances de l'audiovisuel contemporain déterminées par le son : l'effet-clip, l'effet-cirque et la double narration (cette sorte de « mariage blanc »). Le dernier cas a été exploité surtout à travers la voix off qu'a instaurée la figure d'un *pupitre du conférencier*, un personnage hors de la diégèse, placé symboliquement « 'entre' les personnages et nous, fort d'une expérience ou d'un savoir que nous n'avons pas » (JULLIER, 2012, p. 251).

Pourtant les deux premiers cas sont le plus intéressants pour cette recherche : l'*effet-clip*<sup>21</sup> surgit quand la matière de l'expression sonore « instrumentalise » la matière visuelle. Jullier énumère ses caractéristiques :

- Le déploiement d'un objet sonore autonome dont sa structure permet de prévoir son déroulement (l'exemple par excellence, la musique, mais il est possible aussi avec une configuration de bruits ou de paroles).
- Le phagocytage par cet objet sonore des autres éléments de la bande-son (p.e. la suppression des bruits de la rue au profit de la musique).
- Une influence de la structure de l'objet sonore sur le montage image (la synchronisation des points de coupe avec son tempo, p.e.) (Ibid., p. 250).

Le cas contraire est celui de l'« effet-cirque » (déjà noté par Chion), où « l'objet sonore impliqué apparaît alors comme une sorte de supplément un peu artificiel, posé là tout exprès pour accompagner un événement visuel, le souligner (*underscoring*) ou le ponctuer (*mickey-mousing*). Mais la fusion audiovisuelle a échoué : chacun reste dans son coin, bande-son et bande-image » (Ibid. p. 251).

En parlant spécifiquement de l'animation, on contestera cette dernière idée. Néanmoins, l'identification de ces tendances est nécessaire pour sortir d'une description toujours déterminée par la suprématie des configurations narratives ou « diégétiques » des sons et des images. Dans le royaume de l'animation, on trouvera plusieurs exemples de configurations ou ce type de configurations basées plus dans les rythmes ou les contrepoints nous conduisent à des lectures structurales des sons.

---

<sup>21</sup> Voir à ce sujet (JULLIER & PÉQUIGNOT, 2013).

## Rick Altman et le caractère événementiel du son

En plus de théoriser sur les genres cinématographiques et sur la narrativité, Rick Altman est un des pionniers de la théorie du son au cinéma. Son premier ouvrage à ce sujet c'est le numéro monographique de la revue *Yale French Studies* n° 60, « Film/sound », publié en 1980, un des premiers recueils des « sound studies », avec la participation tant des auteurs anglo-saxons, que francophones (ALTMAN, 1980). De ce point de départ, Altman a consacré grand partie de son travail aux sujets relatifs à la dimension sonore du cinéma : l'histoire du cinéma, du muet au parlant, le film musical, ou la théorie générale du son.

Dans ce contexte, l'œuvre la plus intéressante est le deuxième recueil qu'il a dirigé, *Sound Theory, Sound Practice* (ALTMAN, 1992), y compris sept textes de sa main. Ces textes mettent au point plusieurs de ses idées sur le son dans toutes ses dimensions. Je vais m'arrêter sur les trois premiers textes de ce recueil, dont ces idées seront récupérés après par rapport au sujet du son dans l'animation.

La première idée d'Altman qui m'intéresse est la conception du *cinéma comme événement*. L'auteur propose cette idée comme alternative à l'approche plus « textualiste » de ses premiers textes, qui l'ont conduit (lui et d'autres collaborateurs du volume) à omettre plusieurs aspects importants pour la compréhension du phénomène sonore :

*Les sons de films muets sont à peine mentionnés; la technologie du son est presque complètement négligée; Il n'y a pas d'attention aux films non-narratifs, non-longs, ou non-occidentaux. Plus important encore, le son lui-même est le plus souvent traité comme s'il était un véhicule idéal de l'information linguistique ou musicale, reçue par un public anhistorique, dans une situation de visualisation générique, sans aucun objectif spectateuriel* (Ibid., pp. 1-2).

Pour surmonter ces problèmes, Altman propose une nouvelle approche : au lieu de concevoir les films comme des textes, il propose de les considérer comme *événements*. L'auteur offre des modèles graphiques des deux approches, le modèle textuel étant conçu comme un diagramme d'atome, avec un centre, le texte, et des entités conceptuelles gravitant autour de lui : la production, la réception, la culture. L'approche « événementielle », au contraire, est représentée comme un espace tridimensionnel sans gravité où flottent des entités (que représentent les films) avec la forme des tores (ou « beignet », *doughnut-shaped* en

anglais), sans une séparation claire entre l'intérieur et l'extérieur. Dans cet univers, le « texte filmique » n'est plus le centre fermé du modèle atomique, mais

*...il sert de point d'échange entre les deux formes "V", l'une représentant le travail de production, l'autre figurant le processus de réception. Commencé comme un sous-ensemble de la culture en général, une "V" se rétrécit progressivement à mesure que le travail de la production cinématographique suit son cours, d'abord large, avec des idées différentes et des scripts, des postes et des presses, des techniciens et des réécritures, jusqu'à le moment où le travail de production finalement est résolu en un seul produit étroit: le texte. Le processus de réception élargit ensuite à nouveau, pour finalement atteindre le point où il est impossible de distinguer de la culture en général. Dans un monde sans gravité, cependant, ce système de sablier est entièrement réversible. Tout comme la production coule à travers le texte vers la réception, la réception aussi influe régulièrement la production*<sup>22</sup> (Ibid., p. 3).

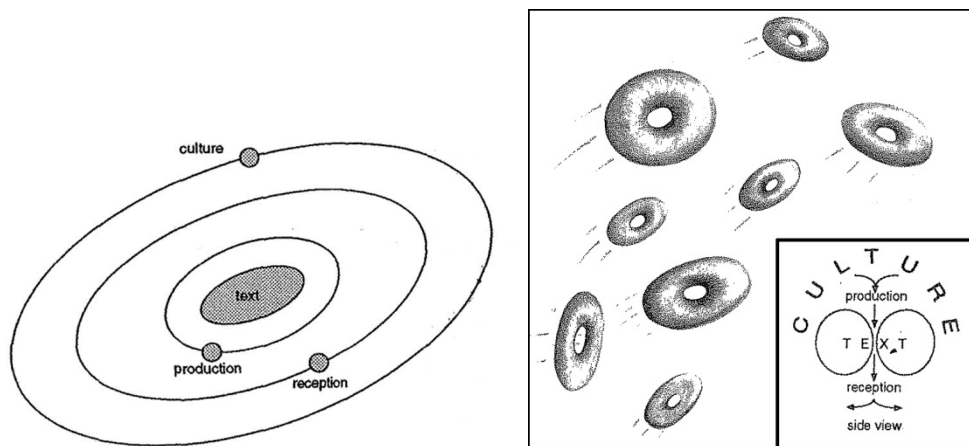


Figure 3 : Des films comme textes aux films comme événements

Source : [ALTMAN, 1992]

<sup>22</sup> “...it serves as a point of interchange between two "V" shapes, one representing the work of production, the other figuring the process of reception. Beginning as a subset of culture at large, one "V" progressively narrows as the work of film production runs its course, first broadly, with diverse ideas and scripts, sets and rushes, technicians and rewrites, until eventually the work of production has been resolved into a single narrow product: the text. The process of reception then broadens out again, eventually reaching the point where it is indistinguishable from the culture in general. In a gravity-free world, however, this hourglass system is entirely reversible. Just as production flows through the text toward reception, so reception regularly influences production”.

L'idée de cette conception « topologique » est d'assumer les films comme entités culturelles ouvertes, où les processus de production et de réception interagissent tout le temps, et où la textualité n'est qu'un point, une étape dans le parcours de production du sens. Ainsi conçu, le cinéma est décrit par Altman à partir de douze « attributs » dynamiques, que je vais réduire à sept, en raison de la redondance de certaines idées :

1. *Multiplicité* : le cinéma a été réduit à l'unité (le film individuel) étant en réalité une entité multiple en sa matérialité (multiples copies pour son exhibition), et dans ses processus de production (multiples rôles et « métiers », multiples créateurs et « partenaires » [*intersection et multidiscursivité*<sup>23</sup>]) et de réception (multiplicité de contextes d'exhibition, de publics, d'individus [*hétérogénéité*]).

2. *Tridimensionnalité* : l'expérience filmique a été réduite aux deux dimensions de l'écran, quand en réalité elle se déploie dans l'espace tridimensionnel de la salle. Cette réduction a conduit à la surestimation de l'image au détriment du son.

3. *Matérialité* : à la différence de la peinture ou la sculpture, dont les critiques et les esthéticiens reconnaissent la matérialité, le film a été abstrait comme un enchaînement d'images « immatérielles » (une chose similaire est arrivée dans le cas de la littérature, laquelle ne se pense pas en termes de support éditorial). Par contre, Altman affirme que « *de la complexité de son financement et de sa production à la diversité de son exhibition, le cinéma doit être considéré en termes des ressources matérielles qu'il engage. Du point de vue*

---

<sup>23</sup> D'après Altman « la complexité de l'expérience cinématographique dérive de la capacité extraordinaire du cinéma de servir comme l'intersection d'une variété de discours, encadrés par divers groupes et adressés aux publics que vont d'individus à l'ensemble de la culture. Partageant le même espace et le temps, ces discours cachent souvent les uns aux autres, dont chaque visualisation d'un film révèle des petits fragments de chaque discours individuel. Un film ne porte pas un message unique, unifié, unilinéaire et univoque. Au lieu, il est plus comme un palimpseste marqué, révélant à divers points des couches discursives diverses, chacune enregistrée à moments différents » (*“the complexity of the cinema experience derives from cinema's extraordinary ability to serve as the intersection of a variety of discourses, framed by diverse groups and addressed to populations varying from single individuals to the entire culture. Sharing the same space and time, these discourses commonly hide one another, with a given film viewing thus successively revealing little bits of each individual discourse. A film does not carry a single message, unified, unilinear, and univocal. Instead, it is more like a scarred palimpsest, at various points revealing diverse discursive layers, each one recorded at a different point in time”*) Ibid., p. 10.

*du son, ce changement est d'une importance capitale, car elle nous évite de tomber dans une conception purement visuelle du cinéma. Comme un produit matériel, le cinéma révèle rapidement l'emplacement et la nature de sa bande son, la technologie utilisée pour le produire, l'appareil nécessaire à sa reproduction, et la relation physique entre les haut-parleurs, les spectateurs et leur environnement physique* » (cet attribut a à voir aussi avec l'hétérogénéité) (Ibid., p. 6)<sup>24</sup>. De cet attribut se dérive aussi un autre : *l'instabilité*, qui éloigne le cinéma de la littérature, dont l'imprimerie a stabilisé et standardisé les textes au détriment de sa matérialité. Le film, au contraire, se rapproche des arts « vivants », même si les technologies d'exhibition (public ou privée) tendent aussi à la standardisation.

4. *Performance* : l'approche textuelle du cinéma a conduit à oublier son caractère performatif, dynamique. Cet attribut est particulièrement important pour l'étape du cinéma muet, où les séances étaient des spectacles divers, mais aussi pour le cinéma en général, toujours impliqué dans des circuits d'exhibition et réception que se déroulent comme une performativité sociale complexe.

5. *Médiation* : les expériences des spectateurs du cinéma ne sont pas isolées de leur expérience avec des médias en général. En ce qui concerne le son, cet attribut est très intéressant parce que, comme Altman le souligne « *le son filmique a été révolutionné plusieurs fois, chaque fois en connexion avec des développements qui affectent des autres industries du divertissement et de la communication, en plus du cinéma. Une fois qu'on reconnaît la nature événementielle du cinéma, le film semble pris dans un réseau complexe de modèles potentiels* » (Ibid., p. 12).

6. *Choix* : concevoir le cinéma comme événement permet d'introduire dans sa description non seulement les discours, produits, séances ou participants effectifs, mais aussi toutes les autres alternatives qui ont été disponibles (et écartées) chaque fois qu'un réalisateur choisit une méthode ou une technologie de production, chaque fois qu'un spectateur choisit un film ou un produit de consommation culturelle déterminé, ou quand un distributeur choisit une technologie de production plutôt que d'autres : « cela sert à mettre en évidence l'importance

---

<sup>24</sup> From the complexity of its financing and production to the diversity of its exhibition, cinema must be considered in terms of the material resources that it engages. From the standpoint of sound, this shift is of capital importance, for it removes cinema from the customary, purely visual definition. As a material product, cinema quickly reveals the location and nature of its sound track, the technology used to produce them, the apparatus necessary for reproduction, and the physical relationship between loudspeakers, spectators, and their physical surroundings.

du cinéma dans le cadre d'un système différentiel, au sens fort de la sémiotique saussurienne: il n'y a pas des termes positifs, seulement des différences » (Ibid., p. 12).

7. *Diffusion (et échange)*: cet attribut a à voir avec le fait que la relation du public avec le cinéma ne s'arrête pas après la séance, mais qu'il continue en se diffusant à travers des autres produits dérivés : les bandes originales, le merchandising, les partitions des chansons des films et, on peut ajouter, tout ce phénomène connu aujourd'hui comme *Trans média*. Cela a à voir aussi avec l'*échange* : l'asymétrie entre le côté de la production et le côté de la réception a été toujours relative, mais à l'actualité cet écart s'est réduit encore plus.

Ainsi Altman ouvre la perspective pour l'analyse du cinéma en général, et du son au cinéma en particulier. Une ouverture qui devra être problématisée, mais qui permet de rendre compte des aspects du cinéma d'animation avant rejetés des analyses.

Dans le deuxième texte d'Altman dans son anthologie, il parle de la *matérialité du son enregistré* (ALTMAN, 1992), pour parler cette fois du *son comme événement*<sup>25</sup>. Ici Altman propose tout un « parcours narratif du son » qui rend compte des phases de la « vie » d'un son, depuis qu'il est produit ou manifesté, jusqu'à ce qu'il soit perçu et interprété.

La première étape de ce « récit » est la *production du son* ; après vient l'*histoire de l'événement sonore*, avec un autre enchaînement de phases « emboîtées » : le parcours de sa production à sa réception (avec toutes ses transformations) ; l'*enregistrement de l'histoire de l'événement sonore* ; la *reproduction* (ou diffusion) de l'*enregistrement de l'histoire de l'événement sonore* ; l'*écoute de la reproduction* (ou diffusion) de l'*enregistrement de l'histoire de l'événement sonore* ; et, finalement, de la *description* de l'expérience d'*écoute de la reproduction* (ou diffusion) de l'*enregistrement de l'histoire de l'événement sonore*. Cette petite joute narrative peut servir comme la synthèse du parcours très complexe de tout procès de description des sons.

## Première synthèse théorique

Ce parcours permet d'identifier quelques aspects théoriques sur lesquels se sont arrêtés les différents auteurs de diverses perspectives théoriques, dans ce qu'on peut appeler la théorie du son au cinéma, et que peuvent être exploités (ou écartés) comme étant utiles pour la description, mais qui doivent être intégrés dans un cadre général. Mon intérêt est d'intégrer

---

<sup>25</sup> Une approche similaire à ce que je vais prendre dans le chapitre 4.

les apports de ces théories dans un cadre *sémiotique*, que j'espère commencer à esquisser ici. On peut donc proposer les suivantes lignes thématiques :

- La description historique de l'avènement du son et de son intégration dans le « langage » ou système expressif du cinéma.
- L'observation des aspects techniques et technologies du son dans l'appareil cinématographique (et ses répercussions dans l'expérience des spectateurs).
- La distinction des matérialités sonores, leurs relations et différences.
- La manipulation discursive des sons dans la production des énoncés audiovisuels (surtout narratifs).
- Les relations du son avec une diégèse (ou « monde possible ») et leurs éléments constitutifs : objets/sources sonores, personnages/sources, espaces et ambiants sonores.
- Les relations et emplacements du son dans une « scène de réception », sa rencontre avec les spectateurs à l'intérieur d'un « dispositif ».
- La caractérisation générique d'un « actant » spectateur (auditeur), selon sa configuration perceptive, cognitive, émotionnelle, corporelle, etc. La diversification des types d'écoutes.

Dans la mesure où la description historique a à voir avec tous les autres aspects et consiste en la description de leurs développements et transformations dans une ligne temporelle, je vais me permettre de la laisser de côté, non pas qu'elle ne soit pas pertinente pour moi –au contraire, la dimension historique sera prise en compte à plusieurs reprises-, mais parce qu'elle est implicite dans chaque ligne d'intérêt.



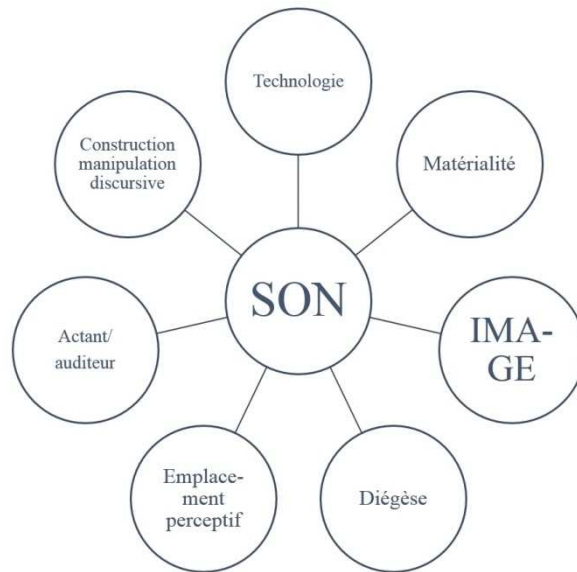


Figure 4 : Relations du son avec les dimensions de l’expérience audiovisuelle

Cela nous donne six types de relation du son avec les diverses dimensions de l’expérience du cinéma. Bien sûr, l’emplacement périphérique de l’image cherche à mettre l’accent sur la dimension sonore, mais cet aspect est aussi central et il est connecté avec les autres, comme le montre la figure XX :

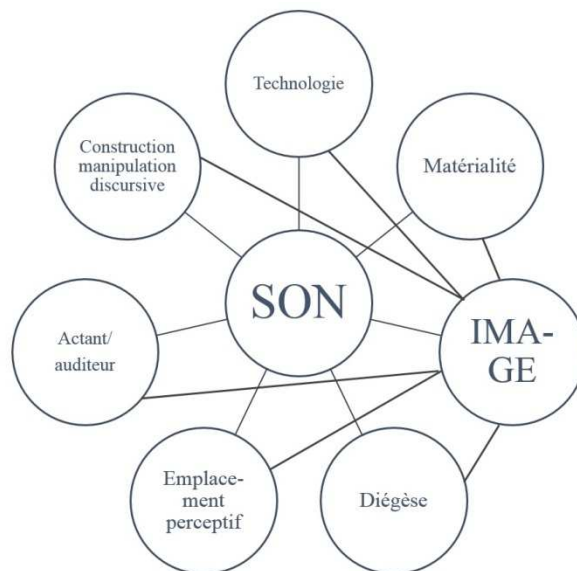


Figure 5 : Relations du son *et de l’image* avec les dimensions de l’expérience audiovisuelle

On peut noter que je n'ai pas mentionné la dimension *textuelle* du son. Cela peut être lu comme un éloignement de la perspective sémiotique qui, traditionnellement, a placé le *texte filmique* comme l'objet de description privilégié (voir Metz, 1971). En effet, l'idée de textualité filmique continue à être centrale aujourd'hui même dans des approches qu'on peut décrire comme *expérientielles*, *phénoménologiques*, ou *pragmatiques* (voir, par exemple, BASSO, 2003), ou dans des approches plus proches des développements de la sémiotique (post)greimassienne (BLANCO, 2003).

Mon omission a été intentionnelle et elle n'est pas due à un rejet de cette notion, toujours nécessaire dans la théorie sémiotique. Pourtant, je veux proposer –en suivant Rick Altman– une *décentralisation* du concept du texte dans la théorie du cinéma, pour souligner ses relations avec d'autres dimensions de l'expérience audiovisuelle. Cela ne veut pas dire qu'on « sorte » de la sémiotique. Au contraire, cela signifie une *expansion*. En suivant Jacques Fontanille, pour qui le texte n'est qu'un *niveau de pertinence* parmi d'autres à l'intérieur d'un *parcours génératif du plan de l'expression*, je voudrais souligner l'importance d'autres niveaux de ce parcours, comme j'espère le montrer dans le chapitre 4 et, surtout, les chevauchements entre ces niveaux. Effectivement, bien que dans les théories narratologiques le texte ou l'énoncé est toujours présent, quand il s'agit du son ou des phénomènes sonores il est moins habituel de trouver des expressions comme « texte sonore », tandis que parler de « texte visuel » est plus courant. Cela peut être dû à l'« essentialisme » dénoncé par Metz, mais aussi au caractère éphémère et « événementiel » des sons, dont on va parler plus loin. En ce sens, mon approche cherche à se solidariser avec ce caractère processuel, dynamique de la sonorité, en l'étendant aux autres dimensions du cinéma. Ainsi, parler d'*expérience audiovisuelle*, comprise comme un ensemble de pratiques circonscrites dans une scène sémiotique, permettra de revenir sur la textualité dans son caractère processuel, même si on l'aborde à partir de sa « solidification » dans un texte fini et « fermé ».

La textualité sera donc abordée comme un procès transversal entre les dimensions proposées avant, et selon des emphases différentes, on la trouvera entre la technologie, la matérialité et la construction discursive, par exemple (du côté des « énonciateurs/producteurs »), ou entre l'auditeur, l'emplacement perceptif et la diégèse (du côté du spectateur du cinéma narratif). Je mentionne ces deux cas parmi plusieurs autres parce que dans le chapitre trois, je vais proposer d'encadrer ces questions sous l'approche de la théorie de l'énonciation (où je reviendrai sur des aspects théoriques à peine ébauchés dans ce

chapitre), par rapport à la théorie du dispositif, présente dans ce parcours initial à plusieurs reprises.

Dans le prochain chapitre j'offre une révision de la littérature sur le son dans le cinéma d'animation, pour tester s'il est possible tout simplement d'appliquer les concepts et les outils de la théorie générale du cinéma, ou s'il y a besoin de développer d'autres catégories pour rendre compte de ce type d'expression audiovisuelle.

## Chapitre 2. Le rôle du son dans les études sur l’animation

Dans un ouvrage récent, Karen Beckman évoque la précarité du traitement de l’animation dans la théorie filmique :

*Une partie de la nature fragmentaire de l'engagement des théoriciens du cinéma à l'égard de l'animation peut provenir du fait que le terme signifie de façons trop différentes. Dans divers moments, le terme devient synonyme de toute une série de concepts beaucoup plus spécifiques, y compris le mouvement, la vie elle-même, la qualité de vivant (qui n'implique pas nécessairement le mouvement) l'esprit, le chromatisme, les processus filmiques cadre-parcadre ou les processus filmiques de cadre variable, le cinéma numérique, ainsi qu'une gamme de supports mobilisés qui apparaissent dans les films d'animation, y compris la sculpture, le dessin, le collage, la peinture, et les marionnettes. Ces termes divergents ne vont pas toujours facilement les uns avec les autres, et bien que les tensions entre eux soient importantes et intéressantes, elles n'ont pas été explorées pleinement comme elles le pourraient, en partie parce que nos paradigmes critiques peuvent avoir forclous ces lignes d'enquête. (BECKMAN, 2014, p. 1)<sup>26</sup>*

D’où le premier problème que nous devons affronter eu égard à ce type de cinéma : la définition elle-même, et les critères pour une telle définition. Comme l’a bien remarqué Suzanne Buchan,

*Les études sur l'animation (animation studies) ont besoin d'un langage qui puisse être spécifiquement utilisé dans les écrits critiques et théoriques sur le film d'animation. Les critiques et les chercheurs sont en train de développer et de définir ce qui a été la condition*

---

<sup>26</sup> “Part of the fragmentary nature of film theorists' engagement with the term may stem from the fact that animation signifies in so many different ways. At different moments, it becomes synonymous with a whole range of much more specific terms and concepts, including movement, life itself, a quality of liveness (that doesn't necessarily involve movement) spirit, nonwhiteness, frame by frame filmmaking processes, variable filmmaking processes, and digital cinema, as well as a range of mobilized media that appear within animated films, including sculpture, drawing, collage, painting, and puppetry. These divergent terms do not always sit easily with each other, and though the tensions among them are important and interesting, they have not been explored as fully as they might, in part because our critical paradigms may have foreclosed such lines of inquiry”.

définie par le filmologue Etienne Souriau pour avoir un « langage bien fait », indispensable à toute discipline scientifique. (BUCHAN, 2006, p. vii)<sup>27</sup>

Et cela est encore plus vrai lorsque l'on parle de la dimension sonore dans l'animation. En tout cas, les approches commencent toujours par l'identification des techniques et des modes de production qui rendent difficile une définition singulière : les dessins animés, l'animation de marionnettes, de pâte à modeler, de papier découpé ou de l'écran d'épingles, des techniques traditionnelles, et des différentes techniques de l'image de synthèse (vectorielle, polygonale, etc.), et plus généralement de l'animation par ordinateur. De ce point de vue, les premières ont en commun la même modalité de production : l'enregistrement photographique image par image, tandis que l'animation numérique produit ses images d'après des logiciels et interfaces, parfois avec des dessins traditionnels comme référence, quelquefois d'après des modèles purement abstraits d'interfaces graphiques<sup>28</sup>.

Néanmoins, la projection unifie ces différentes formes de production dans un même flux d'images que le spectateur perçoit comme mouvement. Ainsi toutes ces formes d'animation s'opposent à l'image-mouvement provenant de la capture directe et « courante » des objets en mouvement réel (ou avec la potentialité de mouvement). Une autre façon de le dire, c'est que le cinéma *live-action* est un flux continu dès la capture, tandis que le cinéma d'animation ne devient flux qu'aux dernières étapes du processus de production (et cela n'est pas tout à fait vrai dans l'animation par ordinateur, qui permet de simuler le mouvement au cours de la même production)<sup>29</sup>.

---

<sup>27</sup> « Animation studies needs a language that can be specifically used in critical and theoretical writings on animation film. Critics and scholars are developing and defining filmologist Etienne Souriau's prerequisite of a "well-made language", essential to any scientific discipline ».

<sup>28</sup> Pour une description des techniques de l'animation, voir LUCENA JÚNIOR (2002) et WELLS (2006).

<sup>29</sup> La définition de l'animation donnée par Norman McLaren est très célèbre: « animation is not the art of drawings-that-move, but rather the art of movements-that-are-drawn. What happens between each frame is more important than what happens on each frame. Therefore, animation is the art of manipulating the invisible interstices between frames » (L'animation n'est pas l'art des dessins qui bougent mais l'art des mouvements dessinés. Ce qui se passe entre chaque photogramme est beaucoup plus important que ce qu'il y a sur chaque photogramme. L'animation est par conséquent l'art de

Mais cette différence n'est pas encore précise, car le spectateur n'est pas toujours au courant des processus de production, mais il peut presque toujours reconnaître la différence entre ces deux tendances de cinéma, même dans des productions hyperréalistes comme *La Légende de Beowulf*, de Robert Zemeckis (2007). Et le problème n'est pas seulement le degré de réalisme de l'image, car il y a des films d'animation réalisés avec des personnes et des objets réels, comme les films de Norman McLaren ou de Jan Švankmajer, faites avec la technique connue sous le nom de « pixilation ». Dans ces cas, le spectateur peut reconnaître que ce sont des films d'animation par l'étrangeté de leurs mouvements.

Du point de vue de la bande-image, dans le cinéma d'animation il est alors nécessaire de différencier le réalisme de l'image elle-même, et le réalisme du mouvement lui-même. Ces deux « réalismes » ne coïncident pas toujours. Dans tout ce qui concerne l'image elle-même, Maurin Furniss a proposé une sorte d'« échelle de l'iconicité » qui peut expliquer, dans une certaine mesure, les différences dans l'apparence visuelle des différentes techniques et styles d'animation (Furniss, 2007, pp. 66-80). Mais dans tout ce qui concerne le mouvement lui-même, il n'y a pas de proposition théorique qui se concentre spécifiquement sur cette question. Il y a, par exemple, des animations qui ont une image avec un réalisme visuel évident, mais une impression de mouvement complètement bizarre ; c'est le cas de la « pixilation », où on retrouve l'apparence de la photographie, mais dont le mouvement nous met en évidence son caractère d'artifice. Il y a, d'autre part, des animations avec une apparence caricaturale, mais avec des mouvements fluides et vraisemblables, comme des animations de Pixar et des autres grands studios d'animation de Hollywood. Il y a aussi des animations picturales ou plastiques, dont le mouvement et l'image ont presque toujours un caractère étrange qui nous rappelle leur artificialité. En tout cas, même l'animation la plus réaliste manque toujours de quelque chose pour accomplir la réalité de la cinématographie régulière. Bien sûr, normalement, ce n'est pas son but. Au contraire, l'animation est attractive précisément par sa différenciation avec l'image « réelle »<sup>30</sup>.

---

manipuler les interstices invisibles qui sont entre les photogrammes). Cité par (SOLOMON, 1987, p. 11). Cette définition a la vertu de ne pas exclure le son, qui aussi « passe entre les photogrammes ».

<sup>30</sup> Dans le chapitre 3 je donnerai une définition plus élaborée de la spécificité de l'animation avec le concept de *dispositif animatique*.

## Le son et la transformation de la réalité

Néanmoins, il existe une autre dimension de l'audiovisuel animé qui participe aussi de la distinction entre le réalisme et « l'expressionnisme » : la bande sonore. À ce sujet, il existe une similitude avec l'image visuelle de l'animation : le caractère artificiel de sa production. Comme souligne Rebecca Coyle :

*Comme un texte audiovisuel qui est complètement construit (c'est-à-dire, en plus des emplacements visuels construits, il n'y a pas des sons enregistrés dans l'emplacement, comme dans le cinéma live-action), l'animation exige de la musique et les films sont d'habitude musicalement saturés (COYLE, 2010)<sup>31</sup>.*

Malgré cette différence, les spécialistes utilisent les mêmes catégories pour parler du son dans l'audiovisuel animé : la distinction traditionnelle entre musique, parole (ou voix) et bruits (quelques-uns distinguent aussi les « sons d'ambiance » et les effets de bruit). Maureen Furniss, par exemple, mentionne trois propriétés acoustiques du son cinématographique : le volume (loudness), le ton (pitch) et le timbre. Selon Furniss :

*Normalement les sons jugés les plus importants sont faits pour être les plus forts; dans la plupart des cas, le dialogue est le plus fort dans la hiérarchie afin de fournir efficacement l'information narrative, les effets et la musique enregistrée étant à un niveau plus faible, comme des éléments de support. Cependant, en changeant la hiérarchie, peut-être par l'augmentation de la musique ou par l'inclusion d'effets sonores qui sont si forts que le dialogue ne peut être entendu, on peut manipuler la perception de l'auditeur sur ce qui se passe dans une scène. (Furniss, 2007, p. 83)<sup>32</sup>*

---

<sup>31</sup> Toutes les citations provenant de l'anglais (et d'autres langues) sont traduites par moi. Je mets toujours le texte original en pied de page: "As audio-visual text that is entirely constructed (that is, in addition to animation's constructed visual locations, there are no existing on-location recorded sounds as in live action), animation is demanding of music and the films are often musically saturated".

<sup>32</sup> "Typically, the sounds deemed most important are made to be the loudest; in most cases, the dialogue is loudest in the hierarchy in order to effectively deliver narrative information, with effects and music recorded at a lower level as supporting elements. However, by changing the hierarchy, perhaps by swelling the music or by including sound effects that are so loud that the dialogue cannot be heard, one can manipulate the listener's perception of what is occurring within a scene".

Néanmoins Furniss ne dit rien de plus du volume ou d'autres propriétés et continue avec les trois aspects mentionnés. Elle commence par les dialogues, et nous raconte comment dans l'industrie américaine ceux-ci étaient enregistrés avant la prise de vues, afin de permettre la synchronisation précise entre les mouvements des lèvres (des personnages dessinés) et les paroles. Un autre avantage de cette pratique est la possibilité de profiter de la voix des acteurs et de ses modulations pour inspirer le design du personnage, sa motricité et les rythmes dramatiques. Mais les voix ne servent pas seulement au développement des actions, mais aussi à multiplier les nuances de la caractérisation :

*Au sein de la vaste catégorie de l'animation, il y a beaucoup plus de variation dans la gamme de voix utilisées que dans les médias de prise de vues réelles (PVR). Bien qu'il existe certainement des exemples de voix relativement « réalistes » utilisés dans l'animation, surtout dans les œuvres réalisées en dehors de l'arène commerciale, dans une grande mesure le dialogue est parlé d'une manière stylisée. (Furniss, 2007, p. 85)<sup>33</sup>*

Furniss décrit des tendances dans l'utilisation des voix dans l'animation : d'un côté, la tendance à la standardisation, dans laquelle un acteur qui trouve une voix reconnaissable par le public est obligé d'utiliser cette voix pour d'autres personnages, sans les différencier. D'un autre côté, la singularisation, grâce à la participation des acteurs célèbres du « star-system », qui transmettent les caractéristiques de leurs personnalités aux personnages animés. Furniss dit aussi que dans l'animation les voix peuvent être altérées à travers des accélérations, ralentissements ou manipulations de leurs propriétés acoustiques. Finalement, cette auteure conclut que les contraintes du contexte –comme le doublage pour différents publics– déterminent aussi l'interprétation des productions animées.

Eu égard aux effets sonores, Furniss se limite à mentionner quelques exemples, en soulignant que « les effets sonores dans l'animation sont souvent appliqués de façon exagérée, sans congruence réaliste entre une action et le bruit créé à partir d'elle » (Furniss, 2007, p. 87)<sup>34</sup>. Son premier exemple en ce sens est *Gerald McBoing Boing*<sup>35</sup>, qui raconte l'histoire

---

<sup>33</sup> “*Within the broad category of animation, there is much more variation in the range of voices used than in live-action media. Although there are certainly examples of relatively ‘realistic’ voices being used in animation, particularly in work produced outside the commercial arena, to a great extent dialogue is spoken in a stylized way*”.

<sup>34</sup> “*Sound effects in animation often are applied in an exaggerated manner, with no realistic congruity between an action and the noise created from it*”.



d'un garçon né sans voix humaine, mais avec des sons onomatopéiques comme son nom : « Boing ! », et cette anomalie lui ouvre finalement le succès dans le monde de la radio.

Dans le milieu industriel, Furniss signale aussi l'utilisation des effets sonores d'une manière complètement singulière, en mentionnent le cas de Tregweth 'Treg' Brown. Ce sonoriste a travaillé avec le studio Warner, et il est le responsable du populaire coyote « Roadrunner » (le Beep-Beep) , une émission qui a influencé le monde de l'animation. Eu égard à leur travail, Chuck Jones, le directeur de l'émission, a dit :

*Je lui ai dit de chercher des incongruités, et il en a fait. Une fois nous avons un film sur le coyote et le chien de berger et nous avons montré des rochers tombant sur le loup. Au lieu d'utiliser l'effet sonore régulier, il a utilisé le son d'une locomotive - et ça a marché<sup>36</sup>.*

Un bon exemple de sa technique est le court métrage *Zoom and Bored*<sup>37</sup>, dans lequel Brown a utilisé tout type d'effets sonores sauf ceux qu'on attend de chaque objet ou action, en créant un effet de drôlerie dans toute la production. Néanmoins l'incongruité n'a été pas identifiée par le public, pas plus par les spécialistes. Furniss conclut cette partie avec un exemple bien éloigné de l'industrie hollywoodienne : le court métrage allemand *Balance*<sup>38</sup>, des frères Lauenstein, qui a gagné l'Oscar en 1989. Dans ce film, le design sonore est subtil et mesuré, et doté de profondeur et de volume, mais aussi de sens dramatique, avec une mise en scène également minimaliste. Ici, la finalité du son est de créer une atmosphère, construire l'espace et ses tensions. Le son accomplit un rôle diégétique, mais aussi dramatique : les inclinaisons du plan et les mouvements sont les protagonistes, plus que les personnages.

Dans la dernière partie Furniss parle de la musique dans l'audiovisuel animé, en accentuant les usages moins commerciaux des compositions musicales. Au début, elle nous rappelle le film *La joie de vivre*, d'Anthony Gross et Hector Hoppin, avec la musique de Tibor Harsányi<sup>39</sup>, qui l'a écrit avant les images, pour permettre une meilleure intégration avec

---

<sup>35</sup> Pour audio-visionner le film : <https://www.youtube.com/watch?v=BI12ZsTR3Mo>

<sup>36</sup> “*I told him to look for incongruities, and he did. Once we had a coyote and sheep-dog film and we showed boulders falling towards the wolf. Instead of using the regular sound effect, he used the sound of a locomotive – and it worked*”. Cité dans Furniss, 2007, p. 87.

<sup>37</sup> Pour audio-visionner le film : <http://www.b99.tv/video/zoom-bored/>

<sup>38</sup> Pour audio-visionner le film : <https://www.youtube.com/watch?v=7wJj58aLvdQ>

<sup>39</sup> Pour audio-visionner le film : <http://www.youtube.com/watch?v=Y2gxu7mD-5c>

la bande image. La pièce musicale a été publiée avant pour sa présentation en concert. Dans le prologue, Harsányi précise que la structure musicale de son œuvre avait des correspondances avec le langage cinématographique : des changements du point de vue, un mouvement permanent du cadre, etc. Harsányi a essayé de s'éloigner des structures purement « scéniques » des compositions symphoniques traditionnelles.

Furniss poursuit avec le cas de l'animateur écossais Norman McLaren qui s'est intéressé à la musique à travers de toute son œuvre. Particulièrement, McLaren s'est intéressé à la synesthésie, le phénomène du croisement des canaux sensoriels. Dans sa jeunesse, l'animateur faisait des « collections d'odeurs », en essayant de faire de la « musique olfactive ». Plus tard, McLaren a cherché des superpositions entre le son et l'image, ou entre son et couleur, en proposant des correspondances entre le ton sonore (pitch en anglais) et le ton chromatique, mais avec la conscience qu'il manque un fondement objectif dans cette équivalence. Il s'est fondé sur des associations subjectives :

*Les tons faibles se rapportent à la couleur pour les valeurs de luminosité faible, quelles que soient la teinte et la chrominance; les tons élevés ont des valeurs de luminosité haute, quelles que soient la teinte et la chrominance. En d'autres mots, je laisse les bruns, les olives, les rouges, les violets, et les bleus obscurs prédominer sur l'écran lors d'un passage à faible hauteur, et j'utilise les jaunes, les verts, les oranges, les rouges, les violets et bleus pâles lors d'un passage de ton aigu.*

*Le timbre du son peut parfois déterminer laquelle de ces couleurs [est utilisée]. Pour moi, il semble lié à la teinte et peut-être davantage à la saturation ou au chromatisme de la couleur (...) En d'autres termes: moins d'harmonie il y a dans le son, moins j'ai tendance à utiliser de saturation ou de chromatisme; et plus riche et plus strident sera le son, plus intense sera la saturation. (FURNISS, 2007, p. 90)<sup>40</sup>*

---

<sup>40</sup> *“Low pitches I relate to colour for low-light values, irrespective of hue and chroma; high pitches to high-light values, irrespective of hue and chroma. In other words, I let dark browns, olives, reds, purples, blues predominate on the screen during a low-pitch passage, and use yellows, pale greens, pale oranges, very pale reds, purples, and blues during a high-pitched passage.*

*The timbre of the sound might sometimes determine which of these colours [is used]. To me it seems related to hue and possibly more so to the saturation or chroma of the colour (...) In other words: the fewer harmonies there are in the sound, the less saturation or chroma I tend to use; and the richer and more strident the sound the more intense the saturation”.*

Dans le même sens, McLaren propose une équivalence entre l'intensité et la taille, bien que «le degré de stimulation des tympans (gamme dynamique) est en corrélation avec le degré de stimulation de la rétine »<sup>41</sup>. Dans sa recherche McLaren s'est intéressé aussi à la production des sons à partir des formes visuelles dessinées dans la bande optique d'enregistrement sonore. Il a conduit des expériences du « son animé » pour voir comment les éléments sonores peuvent être créés à travers le dessin et la fabrication des bandes sonores synthétiques<sup>42</sup>. Dans son film « Neighbors » (Voisins)<sup>43</sup>, le son synthétique achève la narration en créant un accompagnement antinaturel qui affirme le mouvement mécanisé et artificiel des personnages animés avec la technique de la « pixilation ».

Après son compte-rendu détaillé de l'œuvre de Norman McLaren, Furniss s'attarde aussi sur un cas très célèbre : *Fantasia*, de Disney. Ce film est important pour plusieurs raisons : le couplage entre la musique classique et le dessin animé ; le développement d'un nouveau système de son stéréo *avant la lettre* ; et la diffusion massive de ces expérimentations dans le grand public (avec le temps, parce qu'à son époque le film n'a pas été un grand succès). Le *Fantasound* était un système sonore qui permettrait un son d'ambiance dimensionnel, en ouvrant de nouvelles possibilités au son multipiste au cinéma. Jules Engel et Oskar Fischinger<sup>44</sup>, deux animateurs abstraits indépendants, ont travaillé à ce film, et aussi le compositeur Leopold Stokowsky.

Finalement, Furniss couvre très rapidement l'histoire de la musique dans l'animation commerciale, de Carl Stalling à Danny Elfman, du « *mickey-mousing* »<sup>45</sup> à l'affranchissement de la bande sonore comme produit commercial. Elle décrit aussi comment les nouvelles technologies numériques ont changé la production musicale dans l'audiovisuel animé. Néanmoins, elle finit avec un exemple particulier qui enjambe deux moments de l'histoire de l'animation : le compositeur Amin Bathia, qui travaille déjà à l'ordinateur. Mais Bathia souligne que l'influence de Stalling est inévitable même aujourd'hui. D'après lui

---

<sup>41</sup> “*The amount one stimulates the ear-drums (dynamic-range) correlates with the degree to which one stimulates the retina*”.

<sup>42</sup> On reviendra sur l'œuvre de McLaren à ce sujet dans le chapitre 2.

<sup>43</sup> Pour regarder et écouter le film : [http://www.youtube.com/watch?v=i\\_2\\_VYYI4-Y](http://www.youtube.com/watch?v=i_2_VYYI4-Y)

<sup>44</sup> Dont on va parler dans le chapitre 2.

<sup>45</sup> On s'occupera du *Mickey-mousing* dans le chapitre 2.

... [Carl Stalling] a défini le style musical, trop bien connu, qui consiste à jouer tous les gags visuels avec un effet sonore musical. Que ce soit un groupe de violons plumés en marchant sur la pointe des pieds à travers l'écran en parfaite synchronisation, ou les phrases orchestrales de surprise «oh, non !» qui précédaient chaque explosion des appareils Acme, Carl et ses collègues ont créé un ensemble de règles musicales<sup>46</sup>.

Mais ce qui intéresse Furniss dans l'influence du « style » de Stalling sur la musique de l'animation est le fait qu'elle constitue une partie intégrale de sa conception structurale (*structural design*), et même de la structure du montage et du mouvement général, au moins dans l'animation commerciale (mais souvent aussi dans l'animation indépendante). Jusqu'ici, son approche n'impose pas un point de vue, mais elle profite de toutes les dimensions de l'audiovisuel animé, de ses modes de production à ses contextes de diffusion. Et quoique ses exemples soient très intéressants, elle lui manque une méthode appropriée pour des analyses plus précises et profitables. D'ailleurs, Furniss soutient encore la distinction entre les bruits, la voix et la musique, même si elle commence en citant Scott Curtis, qui propose une très large critique de cette distinction.

### **Scott Curtis et les trois mythes de la bande sonore dans les *Cartoons***

L'article de Scott Curtis, *The sound of the Early Warner Bros. Cartoons* a une approche plus historique que théorique, mais les idées qu'il propose constituent des apports importants pour notre projet (CURTIS, 1992). À partir du cas des cartoons produits par la Warner Bros, dans les années 30, Curtis expose la difficulté de parler du son dans l'animation avec la même terminologie que celle utilisée pour le cinéma en général :

*La terminologie ne correspond pas tout à fait aux phénomènes. Trois schèmes communs –la hiérarchie image/son, la séparation de la piste sonore en dialogue / musique / effets, et la distinction « diégétique/non-diégétique »– sont particulièrement difficiles à manier lorsqu'il s'agit de l'animation, ce qui suggère que nous devrions repenser nos descriptions du son au*

---

<sup>46</sup> “[Carl Stalling] defined the all too well-known musical style of playing every visual gag with a musical sound effect. Whether it be a group of plucked violins tiptoeing across the screen in perfect sync, or the surprise ‘oh no’ orchestral phrases that preceded every exploding Acme device, Carl and his peers created a set of musical rules” Cité dans (Furniss, 2007, p. 94).

*cinéma pour y inclure d'autres types de textes filmiques en plus des longs métrages de prise de vues réelles.* (CURTIS, 1992, p. 195)<sup>47</sup>

Il est curieux que ces trois distinctions théoriques soient formelles, alors que l'approche de Curtis pour résoudre les problèmes qui suivent est de type technologique ou même économique. Cela est dû au fait qu'à l'époque, les contraintes techniques et économiques de la production du cinéma animé ont donné lieu aux déterminations formelles des cartoons. D'un côté, les droits commerciaux sur l'usage de la musique font que les studios profitaient de leurs compositions, de sorte que les pièces musicales jouaient un rôle très important dans le design structural des cartoons. Il est vrai que les deux grands studios hollywoodiens avaient deux séries chacun : Disney avait la série de Mickey Mouse et les *Silly Symphonies*, tandis que Warner Bros avait les *Looney Tunes* et les *Merry Melodies*. D'après les historiens, dans une des séries la musique devait suivre les actions dessinées, et dans l'autre c'était le contraire : les actions devaient s'accommoder à la musique. Mais Curtis soutient que dans toutes ces séries, les réalisateurs ont développé des techniques très sophistiquées pour synchroniser les dessins et la musique, de sorte que cette dernière était toujours fondamentale pour la conception structurelle de l'animation. En outre, les technologies d'enregistrement sonore n'avaient pas une grande qualité et les micros étaient calibrés de façon irrégulière, donc la musique finissait par remplacer les effets sonores.

Concernant la deuxième distinction dans l'animation: diégétique/non diégétique, Curtis est catégorique :

*La distinction «diégétique / non diégétique» ne s'applique même pas à la bande-image du dessin animé si l'on considère que l'espace et le temps allongés dans les scènes de poursuite sont plus étroitement liés au rythme musical qu'à un espace unifié fictivement. Ainsi, la traditionnelle catégorisation «diégétique / non diégétique», lorsqu'elle est appliquée même au dessin animé le plus ordinaire, est rendue pratiquement inutile. La relation entre la musique et l'image dans des dessins animés peut être mieux décrite par la distinction entre ce qu'on*

---

<sup>47</sup> “The terminology does not quite fit the phenomena. Three common schemata –the image/sound hierarchy, the separation of the sound track into dialogue/music/effects, and the “diegetic/non-diegetic” distinction– are particularly unwieldy when dealing with animation, suggesting that we should reconceptualize our descriptions of film sound to include other types of film texts besides live-action features”.

*pourrait appeler les usages isomorphes et les usages iconiques du son. (CURTIS, 1992, p. 201)<sup>48</sup>*

D'après Curtis, le son « isomorphe » est celui qui a la même « forme » (*shape*) que l'image. L'isomorphisme réfère ici au synchronisme précis entre son et image, dans une relation basée sur le rythme tant de l'action que de la musique. Le meilleur exemple de ce type de relation est le « *mickey mousing* »<sup>49</sup>, une pratique très commune dans le cartoon classique, inaugurée par le court métrage *Steamboat Willie*<sup>50</sup> (le premier dessin animé avec un son synchronisé) dans lequel les mouvements des personnages et des objets sont presque complètement coordonnés avec la musique. Néanmoins, il y a parfois des sons « musicaux » qui jouent un rôle très différent, par exemple quand le son d'une cymbale coïncide avec le coup qu'un personnage donne à un autre : « les moments comme ceux-ci, quand l'orchestre produit des « effets sonores » (diégétiques ?), pourraient être appelés 'iconiques' »<sup>51</sup>. Tandis que les relations isomorphes sont fondées sur le rythme, les relations iconiques sont des relations analogiques entre les événements visuels et le timbre, le volume, et le ton du son. D'après Curtis :

*L'avantage des termes « iconique / isomorphique » est qu'ils ne sont pas mutuellement exclusifs, comme les termes « diégétique / non-diégétique », et ils peuvent décrire diverses variables sonores. Dans un cas complètement différent, le film d'Oskar Fischinger Motion Painting No. 1, la ligne de basse du Brandenburg Concerto de Bach s'élargit en même temps que dans l'image les cercles s'élargissent et que leur teinte augmente. Cela semble être un isomorphisme dans la mesure où le rythme de la basse et le mouvement des cercles coïncident, mais il est aussi iconique dans la mesure où le changement des couleurs correspond à l'approfondissement de la hauteur. Si les relations isomorphes se réfèrent à celles régies par*

---

<sup>48</sup> “The «diegetic/non diegetic» distinction does not even apply to the image track of the cartoon if we consider that elongated space and time in chase scenes are more closely related to musical rhythm than to a fictionally unified space. Thus, the traditional « diegetic/non diegetic » categorization, when applied to even the most ordinary cartoon, is rendered practically useless. The relationship between music and image in animated cartoons can be better described by distinguishing between what might be called isomorphic and iconic uses of sound”.

<sup>49</sup> On s'occupera du *Mickey mousing* dans le chapitre 2.

<sup>50</sup> Pour audio-visionner le film : <http://www.youtube.com/watch?v=BBgghnQF6E4>

<sup>51</sup> “Moments such as these, when the orchestra provides the (diegetic?) ‘sound effects’, might be called ‘iconic’”.

*le rythme et le mouvement, les relations iconiques appartiennent à des relations analogiques entre les événements visuels et le timbre, le volume, la hauteur et le ton du son conjoint.* (CURTIS, 1992, p. 202)<sup>52</sup>

D'après Curtis, tandis que la relation entre les comédiens et leurs voix enregistrées est indicielle, les voix des cartoons (jouées aussi par des comédiens, mais exagérées ou déformées pour sembler « drôles ») ont une relation *iconique* avec les personnages animés, dans la mesure où l'exagération de la voix s'accorde avec le caractère « malléable » des dessins. Ainsi, Curtis remarque la distinction entre l'animation et les autres formes audiovisuelles et nous invite à étudier la première avec des instruments théoriques plus aiguisés que ceux qu'on utilise pour les usages traditionnels du son.

### **Dessinés pour sonner (Drawn to Sound) : la « sonicité » de Rebecca Coyle**

Il n'y a pas beaucoup d'œuvres consacrées exclusivement au sujet spécifique de cette thèse. Une des exceptions est le livre édité par Rebecca Coyle : *Drawn to Sound. Animation Film Music and Sonicity*, dans lequel cette chercheuse australienne assemble des articles concentrés sur un même but : montrer comment le son joue des rôles particuliers dans l'audiovisuel animé. En plus de participer avec deux articles dans cette compilation, Coyle fait l'introduction, qui fonctionne comme un « état de l'art » sur le sujet. Pour cadrer son approche, Rebecca Coyle essaie de proposer un nouveau terme, la *sonicité* :

*J'ai emprunté le terme « sonicité » (sonicity<sup>53</sup>) à la science, où il est fait référence à la transmission du pouvoir par forces et mouvements périodiques. Ce sens a été mobilisé par des*

---

<sup>52</sup> “The advantage of the terms «isomorphic/iconic» is that they are not mutually exclusive, as are ‘diegetic/non-diegetic,’ and they can describe multiple sound variables. In another, completely different example, Oskar Fischinger’s Motion Painting No. 1, the bass line of Bach’s Brandenburg Concerto broadens as the circles in the image broaden and deepen in hue. This seems to be isomorphic in that the rhythm of bass and the movement of the circles coincide, but it is also iconic in that the change of colors matches the deepening pitch. If isomorphic relations refer to those governed by rhythm and movement, then iconic relations pertain to analogous relationships between visual events and the timbre, volume, pitch, and tone of the accompanying sound”.

<sup>53</sup> La sonicité est une invention de George Constantinesco, ingénieur roumain qui a travaillé dans le domaine de la mécanique des solides. La sonicité décrit la transmission de l'énergie par des vibrations : « le principe des procédés Constantinesco est le suivant : l'énergie est transmise au moyen de vibrations se propageant à travers un milieu liquide, de même que les ondulations d'un courant



*chercheurs et des artistes associés au World Soundscape Project (voir Stanza et Kandola en ligne<sup>54</sup>). Comme le but ultérieur de ce projet est de «trouver des solutions pour un paysage sonore écologiquement équilibré où la relation entre la communauté humaine et son environnement sonore soient en harmonie», il est pertinent pour la discussion sur le son dans l'animation, dans lequel la création idéalisée d'un monde, ses habitants et leurs activités est un projet partagé du son et de l'image. (COYLE, 2010, p. 3)<sup>55</sup>*

Coyle souligne que dans le domaine des études en animation, les chercheurs parlent presque toujours sur l'image, en omettant le son. Celui-ci est pourtant une dimension qui permet de concevoir des aspects inédits sur le mouvement dans l'espace et le temps :

*Le son ne peut être congelé de la même manière que les images peuvent être présentées sur une page, en dépit des meilleurs efforts de musicologues pour capturer des éléments dynamiques à travers la notation de mélodies et d'arrangements. Le son est mouvement constant (...) Le son se déplace également dans l'espace stéréophonique et à travers des transitions entre les scènes, les cadres et les juxtapositions. Il fournit ses propres*

---

électrique se propagent le long d'un conducteur métallique. Les ondes ainsi employées sont de même nature que les ondes sonores, et c'est précisément en étudiant l'harmonie musicale et ses lois, que M. Constantinesco a été conduit aux applications mécaniques que nous allons exposer. De là le nom de sonicité qu'il a imaginé pour grouper sous une même appellation les phénomènes dont il a étudié les lois et recherché les applications ». Pour voir l'explication complète du principe de la sonicité : <http://www.gloubik.info/sciences/spip.php?article579>

<sup>54</sup> Il ne se trouve plus à l'adresse électronique donnée par l'auteure. L'histoire du *World Soundscape Project* se trouve ici : <https://www.sfu.ca/~truax/wsp.html> Le leader de ce projet est R. Murray Schaefer, créateur du concept de *paysage sonore* et auteur du livre avec le même titre (SCHAFER, 2010).

<sup>55</sup> *“I have borrowed the term ‘sonicity’ from science, where it refers to the transmission of power by periodical forces and movements. This sense has been mobilized by researchers and artists aligned with the World Soundscape Project (see Stanza and Kandola online). As the project’s ultimate aim is “to find solutions for an ecologically balanced soundscape where the relationship between the human community and its sonic environment is in harmony”, it is apposite for the discussion of sound in animation, in which the idealized creation of a world, its inhabitants and their activities is a shared project of sound and image”.*



*enchaînements, préfigurant et propulsant la narration, les «points de drôlerie» pour mettre en évidence la main de l'animateur, sa «personnalité». (COYLE, 2010, p. 4)<sup>56</sup>*

En plus d'offrir des arguments sur l'importance du son dans l'animation (dialogues, musique, bruits), Rebecca Coyle inclut dans son introduction tous les facteurs qui permettent d'expliquer l'absence de réflexions sur ce sujet : facteurs industriels, économiques, esthétiques et culturels, parmi d'autres. Elle offre aussi un panorama général de la littérature sur le son dans l'animation.

La première grande référence est le livre de Giannalberto Bendazzi, Manuele Cecconello et Guido Michelone, *Coloriture : Voci, rumori, musiche nel cinema d'animazione* (BENDAZZI, CECCONELLO, & MICHELONE, 1995). Rebecca Coyle fait une synthèse superficielle de cette œuvre –qui à notre avis mérite un commentaire plus détaillé que nous ferons ci-dessous–, en disant que les auteurs proposent une série de réflexions sur la musique dans l'animation, avec des profils de réalisateurs et compositeurs et des analyses de productions concrètes. Et comme appendice, il y a dans le livre aussi quelques renseignements sur le doublage des voix, les effets sonores et la synesthésie.

Du côté américain, Coyle mentionne l'anthologie *The Cartoon Music Book*, dans laquelle Daniel Goldmark et Yuval Taylor rassemblent des textes et réflexions de plusieurs théoriciens, musiques et réalisateurs sur les rôles de la musique dans le cartoon américain (GOLDMARK & TAYLOR, 2002). Bien que cette œuvre ne soit pas proprement théorique (ou plus précisément : académique), elle offre des renseignements et des analyses sur les créateurs. Dans cette même tendance, Coyle parle de Richard Davis, qui consacre un chapitre de son livre *Complete Guide To film scoring* (DAVIS, 1999), à montrer les différences entre les approches musicales à la télévision, dans les longs-métrages, les films musicaux et les œuvres dramatiques. Dans une autre œuvre plus théorique David Goldmark propose un parcours historique des grands studios américains et montre les différents chemins pris par la musique des cartoons traditionnels en décrivant avec précision leurs modèles structurels (GOLDMARK, 2005).

---

<sup>56</sup> “*Sound cannot be freeze-framed in the same way that images can be presented on the page, despite the best efforts of musicologists to capture dynamic elements by notating melodies and arrangements. Sound is constant movement (...) Sound also moves around stereo space and transitions across scenes, frames and juxtapositions. It provides its own segues, foreshadowing and momentum for narrative, ‘gag points’ to highlight the animator’s hand, and ‘personality’.*”

Un autre auteur mentionné par Coyle est Robin Beauchamp, qui dans son livre *Designing Sound for Animation* (BEAUCHAMP, 2005) alterne des réflexions techniques avec des propositions théoriques, en combinant son expérience comme producteur musical et comme enseignant. Son œuvre comporte tout un chapitre sur la théorie du design sonore, mais aussi des idées générales sur la perception du son et de l'image. Après cela, Beauchamp consacre un chapitre à chacun des éléments sonores : dialogues, musique et effets sonores<sup>57</sup>.

Rebecca Coyle mentionne aussi quelques articles isolés loin de la sphère anglo-américaine, particulièrement des auteurs australiens : Philip Brophy et ses contributions dans les deux livres sur le sujet édités par Alan Cholodenko : *The Illusion of life I et II* (BROPHY, 1991) ; et Michael Hill, qui a écrit l'article *Life in the Brush : the Orchestration of Nature in Australian Animated Feature Films*, où il décrit la « musicalité » de quelques films australiens d'animation. Coyle mentionne aussi quelques œuvres sur Walt Disney et sa compagnie discographique, *Walt Disney Records*. Rebecca Coyle conclut son introduction en soulignant la prolifération des publications et des journaux spécialisés dans la musique et le son au cinéma et dans l'audiovisuel, mais aussi dans l'animation. Néanmoins, à son avis les premières manquent d'une approche intégrale pour l'audiovisuel et les dernières ne s'arrêtent presque jamais sur le son.

Pourtant, l'œuvre éditée par Rebecca Coyle se concentre elle-même plus sur la musique que sur la totalité de la bande-son, et c'est seulement dans la troisième partie – précisément la partie que reprend le titre du livre, *musique et sonicité* –, qu'elle fait son retour comme protagoniste. Le premier des articles de cette partie est *Sonic Nostalgia and Les Triplettes de Belleville* (GOLDMARK, 2010), dans lequel Daniel Goldmark réfléchit sur l'œuvre de Sylvain Chomet. Goldmark commence en se demandant pourquoi Chomet a choisi ne pas utiliser les dialogues dans son œuvre, et c'est l'auteur lui-même qui offre la réponse :

*Je pense aussi qu'une animation sans les contraintes des paroles est plus forte. Si tu dois adapter tous les mots, tout le mouvement gestuel tourne autour de la bouche. Sans elles, tu es beaucoup plus libre de créer une véritable animation, à travers l'animation elle-même*<sup>58</sup>.

---

<sup>57</sup> Ci-dessous nous faisons un commentaire plus détaillé de cette œuvre.

<sup>58</sup> “*I also think that an animation without the constraints of spoken words is stronger. If you have to fit everything to the words, all the gestural movement revolves around the mouth. Without it, you are much freer to create true animation, to through animation itself*”. (GOLDMARK, 2010, p. 142).

Goldmark rapproche cette idée des premiers cartoons sonores américains, où les dialogues n'étaient plus le centre de l'animation, mais les gags visuels et la musique.

Goldmark s'arrête sur le premier film de Chomet, *La Vieille Dame et les pigeons* (1998)<sup>59</sup>, pour montrer le rôle du son dans ce court métrage :

*Les effets sonores supportent le spectacle ici, cependant, en particulier les sons de tapotement, battement, roucoulement des pigeons, et les sons évocateurs de la nourriture qui est déchirée, mâchée et avalée. Chomet veut que nous sympathisions avec cet homme, et peut-être même que nous ressentions autant la faim que lui*<sup>60</sup>.

Même les voix des touristes au début du film fonctionnent comme bruitage. Il y a aussi de la musique, et le rythme de l'animation est synchronisé avec elle, selon le moment dramatique de l'histoire, quand le gendarme est sur le point d'être dévoré. Goldmark dit que quand le dialogue manque, la bande-son est élargie et permet un design sonore plus complexe, et une meilleure attention à la narration visuelle, en lui permettant de profiter la totalité de l'espace sonore que le film offre.

Ensuite Goldmark, se consacre à l'analyse du film *Les Triplettes de Belleville*, en affirmant qu'il tourne autour de la nostalgie. Dès le début du film, avec ce petit « film dedans le film », le vidéo-clip des triplettes dans leur jeunesse, en tonalité sépia et avec des ratures des vieilles productions et la musique typique des années 30, jusqu'aux transitions des « paysages sonores » dans la ville où vivent les personnages, sont décrits avec précision dans ce long passage qui nous permettons de reproduire en totalité :

*Les sons de l'industrie marquent discrètement l'évolution que le temps a forgée. La séquence se passe si vite, soulagée par la musique, que nous ne pourrions pas même remarquer cette touche subtile dans le montage sonore: le second coup en un grand plan unique, avec le bruit des hélices qui tournent, tandis que le troisième coup de feu correspond à deux avions à réaction. La transition sonore du moteur au turboréacteur indique non seulement des progrès, mais aussi le profond sentiment de perte qui imprègne ce film. Comme le bourdonnement de*

---

<sup>59</sup> Pour audio-visionner le film: <https://www.youtube.com/watch?v=MqhVcRrauY>

<sup>60</sup> “The sound effects run the show here, however, particularly the sounds of the pigeons tapping, flapping, cooing, and the evocative sounds of food being torn, chewed and swallowed. Chomet wants us to sympathize with this man, and perhaps even feel as hungry as he is” (GOLDMARK, 1010, p. 144).

*l'hélice se fane dans le passé, un symbole de technologie dépassée, le rugissement du moteur du jet signifie un cadeau indésirable, ou que le temps passe trop vite. Le temps pour ces personnes est ce qu'on doit chérir, pas ce qu'on quantifie. La prédiction de Schafer (d'une «surdit  universelle» avec la pollution sonore en cours) trouve une nouvelle application ici quand Chomet opte pour la mutit  universelle de ses personnages. Par ces moyens, il oblige le public   entendre le bruit (en vertu de l'omission du dialogue) dans les divers contextes g ographiques (la campagne vs la ville, les montagnes vs l'oc an, l'Europe vs l'Am rique du Nord)<sup>61</sup>.*

Cette subtile analyse nous permet de concevoir le type de r les que le son peut jouer dans la construction narrative des dessins anim s. Goldmark montre aussi comment le son permet de construire les points de vue (ou les points d' coute ?) de la narration, m me celui d'un chien dans ce cas. Cet article est presque le seul qui offre une d monstration de ce que peut faire la *sonicit * dans l'audiovisuel anim .

Les deux autres articles de cette partie s'occupent plus de la musique, quoiqu'ils mentionnent aussi l'importance des effets sonores, et la difficult  de tracer une borne entre ces deux composants de la bande-son. Dans *Resilient Appliances : sound, Image and Narrative in The Brave Little Toaster*, Jon Fitzgerald et Philip Hayward essayent de montrer comment cette production, rejet e par les Studios Disney, a ouvert de nouveaux parcours au design sonore et musical du film anim , particuli rement, avec le choix des voix et la construction narrative   travers des chansons originales (FIZGERALD & HAYWARD, 2010, pp. 160-173).

---

<sup>61</sup> *The sounds of industry discreetly mark the changes time has wrought. The sequence goes by so quickly, eased along by the music, that we might not even notice such a subtle touch in the sound editing: the second shot has a single large plane, complete with the sound of the propellers turning, while the third shot has two jet airplanes. The aural transition from prop motor to jet engine indicates not only progress, but also the profound sense of loss that pervades this film. As the buzzing of the propeller fades into past, a symbol of outdated technology, the jet-engine roar signifies an unwelcome present, where time passes too quickly. Time for these people is for treasuring, not quantifying. Schafer's prediction (of a "universal deafness" with ongoing noise pollution) finds a reapplication here as Chomet opts for universal muteness for his characters. By these means he compels the audience to hear (by virtue of omitting dialogue) the noise as in the various geographical settings (country versus city, mountains versus ocean, Europe versus North America) Ibid. p. 148.*

Pour sa part, Kentaro Imada démontre comment le design sonore et musical dans les films et émissions de télévision au Japon mélange une tradition qui remonte au théâtre Kabuki –*hayashi*, l’accompagnement sonore et dramatique dans la salle, nommé ensuite *gekiban* au cinéma– avec la tradition occidentale de la musique cinématographique, en analysant une animation spécifique : les séries et films autour du personnage d’Arsène Lupin III, le petit-fils du personnage des romans de Maurice Leblanc. Imada peut prouver à partir de ce cas que, dans l’animation japonaise, la musique –même quand elle est écrite et jouée par de prestigieux musiciens, comme Yuji Ohno, qui a composé une musique *jazz* de style occidental– est subordonnée au travail du *senkyokuka* (le « sélecteur musical », mais qui fonctionne comme directeur sonore en général). En ce sens :

*Les pratiques décrites ci-dessus contrastent avec la perception auteuriste classique des compositeurs de musique artistique occidentale. Dans le monde de l’animation japonaise, plutôt que de créer des mélodies one-of-a-kind pour l’adapter aux séquences d’images particulières, le travail du compositeur est en grande partie de développer un stock de matériau qui peut être utilisé quand nécessaire. Ce n’est pas juste une question d’opportunité économique, mais cela reflète également les racines du gekiban comme une pratique scénique et médiatique japonaise. (IMADA, 2010, p. 184)<sup>62</sup>*

Pourtant, le seul article qui s’arrête vraiment sur le son et les bruits est celui de Goldmark. Ainsi, la *sonicité* de Rebecca Coyle est loin d’être un concept complet et fonctionnel pour parler du son au cinéma d’animation, quoiqu’elle ouvre tout un parcours théorique en attente d’être développé et auquel j’espère contribuer.

### ***Designing Sound for Animation* de Robin Beauchamp : vers une rhétorique du design sonore**

Nous avons déjà parlé du livre de Robin Beauchamp d’après la révision de Rebecca Coyle. Maintenant je voudrais souligner quelques aspects de cette œuvre, faite par un professionnel du design sonore, mais qui offre tout un répertoire d’idées et de concepts

---

<sup>62</sup> “*The practices described above contrast to classic auteurist perception of western art-music composers. In the world of Japanese animation, rather than creating one-of-a-kind melodies to fit particular image sequences, the job of the composer is largely to develop a stockpile of material that can be used as necessary. This is not just a matter of economic expediency, but also reflects the roots of gekiban as a practice in Japanese stage and media productions*”.

importants pour comprendre le « langage du son » dans la tradition conventionnelle américaine. En particulière nous parlerons du chapitre 2 : « Sound design theory », dans lequel Beauchamp propose les assises conceptuelles de l'utilisation narrative du son (Beauchamp, 2005).

Dans l'introduction du chapitre, Beauchamp commence en disant que, d'après beaucoup de réalisateurs, le son contribue à 70% du succès d'un projet (mais, d'après Rebecca Coyle, il exagère un peu). Ensuite, il propose une série de classifications du son, principalement à partir de Michel Chion, mais toujours en soulignant des équivalences avec le langage technique. Ainsi, la musique diégétique s'appelle *source music*, tandis que la musique non diégétique est nommée *under-score*. Les effets sonores liés aux objets dans le cadre sont *hard effects*. La pratique générale qui consiste à ouvrir une clé musicale ou un environnement sonore est nommée *establishing sound* (l'établissement du son). En plus de cette terminologie, Beauchamp reprend les concepts de Chion (adaptation des « écoutes » de Schaeffer), les sons causaux (ou littéraux), sémantiques (ou symboliques) et réduits (ou acousmatiques), pour conclure que les derniers sont les plus importants pour un design « créatif » (tandis que les sons causaux sont utilisés surtout pour produire un effet de réalisme, et les sons sémantiques sont presque uniquement des paroles).

Pour continuer, Beauchamp décrit les différences entre le traitement visuel et le traitement sonore, pour montrer comment ces deux types d'information sont complémentaires : « Le design sonore peut aider le public à percevoir avec précision des images complexes; cependant, il peut aussi faire que des images simples se sentent plus complexes »<sup>63</sup>. Notre champ visuel est limité à 180 degrés, tandis que le son est perçu à 360 degrés ; le traitement de l'information sonore est aussi plus rapide, donc l'écoute nous aide à accentuer l'image et agrandir l'expérience audiovisuelle : « si la conception sonore est efficace, le public se souviendra des détails visuels d'une scène qui n'existait que dans la dimension acoustique » (BEAUCHAMP, 2005, p. 18)<sup>64</sup>.

---

<sup>63</sup> “Sound design can help the audience accurately perceive complex visuals; however, it can also make basic visuals feel more complex”.

<sup>64</sup> “If the sound design is effective, audiences will recall visual details from a scene that only existed sonically”.

Pour continuer, Beauchamp décrit une série de caractéristiques ou fonctions du son dans l'animation, que nous résumons ici :

**Influence du son dans la perception du temps :** pour illustrer cette influence, Beauchamp nous invite à regarder une scène de poursuite sans son : nous découvrons rapidement qu'il y a beaucoup d'images relativement statiques et que les prises en close-up intercalées dans l'action neutralisent l'impression de mouvement quand les sons et la musique nous manquent. Les effets visuels de ralentissement, d'accélération ou de saut dans le temps ont aussi leur contrepartie sonore, qui renforce l'image ou contraste avec elle. Finalement, le son permet d'agrandir les possibilités d'attention des spectateurs :

*Les spectateurs ont des périodes d'attention de plus en plus limitées, ce qui influe sur leur perception du temps en dehors du récit. (...) Quand les critiques parlent de la longueur du film, ils parlent souvent de la qualité du film. Les éléments sonores sont des moyens essentiels de l'extension de l'attention tout en laissant le public avec le sentiment que le temps a été suspendu. (BEAUCHAMP, 2005, p. 19)<sup>65</sup>*

**Influence du son dans la perception de l'espace :** ici la technique est essentielle dans le sens que l'arrivée du son multicanal (le son « surround ») a étendu les frontières de l'expérience audiovisuelle hors de l'écran : « lorsqu'il est utilisé de manière efficace, le son peut améliorer les concepts spatiaux tels que la profondeur, la largeur et la hauteur »<sup>66</sup>. Beauchamp décrit avec une grande précision technique comment la manipulation des caractéristiques du son (volume, ton, fréquence, réverbération, mouvement, direction, etc.) produit des transformations dans notre expérience de l'espace (BEAUCHAMP, 2005, pp. 19 - 20).

**Attirer le public dans le récit :** c'est le rôle de l'*establishing sound*. Comme l'histoire progresse, le son produit l'effet de clôture d'une scène et attire le public vers la prochaine. À la fin du film, la musique et le son produisent l'effet de clôture définitive, en facilitant la transition vers la réalité.

---

<sup>65</sup> *“Audiences have increasingly limited attention spans, a fact that influences their perception of time outside of the narrative. (...) When reviewers talk about the length of the film, they often are talking about the quality of the film. Sound elements are essential means of extending attention spans while leaving the audience with the feeling that time has been suspended”.*

<sup>66</sup> *“When used effectively, sound can enhance spatial concepts such as depth, width and height”.*



**Le son comme développeur des personnages :** dans le cas des personnages humains, Beauchamp parle surtout de la voix et de ses caractéristiques sonores (ton, volume) pour les individualiser, mais aussi des *leitmotifs* musicaux, un héritage de l'Opéra très commun dans l'animation. Quant aux personnages non humains, Beauchamp évoque leur anthropomorphisation, car dans l'animation tous les objets ont la potentialité de déployer des caractéristiques humaines<sup>67</sup>. C'est un défi pour les concepteurs sonores qui doivent développer les sons les plus plausibles pour ces objets.

**La plausibilité et la réalité :** d'après Beauchamp, dans l'animation ne s'agit pas du réalisme, mais de plausibilité, à travers l'analogie, le contraste, l'exagération : « ironiquement [dans l'animation] le son provenant vraiment d'une source est souvent moins convaincante ou divertissante que la substitution exagérée » (BEAUCHAMP, 2005, pp. 21-22)<sup>68</sup>.

**Le son métaphorique :** ce type de sons permet d'éviter des interprétations littérales des animations, et d'augmenter le caractère dramatique des situations et des personnages. Mais le son métaphorique permet aussi souvent la caractérisation effective d'un objet ou personnage, bien que beaucoup de ces objets n'existent pas dans la réalité et partant n'ont pas un son « littéral ». Quant au caractère dramatique, Beauchamp donne l'exemple d'un clignement : dans la réalité il ne sonne pas, mais avec un effet sonore dans une animation il se transforme en un événement.

**Le son hors cadre :** le son nous donne la possibilité de mettre en scène n'importe quelle action, qu'elle soit très chère, très polémique ou très longue pour le seul développement visuel. N'oublions pas que Beauchamp parle du point de vue d'un réalisateur.

**Tension et détente :** Beauchamp décrit comment les variations dans le ton de la musique, la voix et les sons produisent des effets de tensions ou de détente chez les spectateurs en faisant de la bande-son le responsable en grande mesure de leurs réponses émotionnelles et, même, physiologiques, comme l'augmentation des battements du cœur.

---

<sup>67</sup> Et pas seulement des caractéristiques humaines: dans l'animation un objet peut être transformé en un animal, une force de la nature ou n'importe quelle entité dotée d'un mouvement propre.

<sup>68</sup> “Ironically, [in animation] the real sound of a source is often less convincing or entertaining than the exaggerated substitution”.



**Adoucir le montage visuel :** à la différence de notre expérience en état de veille, le montage filmique nous offre un flux discontinu d'images, avec des coupes, des ellipses et des interruptions. D'après Beauchamp, le son fournit la linéarité qui adoucit les effets de la progression non linéaire des images. L'auteur souligne que les sons sont presque toujours linéaires, sans découpages, et les images « parasitent » cette linéarité<sup>69</sup>.

**Développement de la continuité :** ici continuité signifie unité. Un film doit sonner toujours avec un même « style », ou la même « tonalité ». C'est pour cela que la musique d'une production est faite presque toujours par un seul compositeur.

**Interprétation du montage visuel et des conventions filmiques :** ici Beauchamp propose tout un répertoire de conventions visuelles et de leurs corrélats sonores : mouvements de la caméra, formes de découpage, types de plans, etc. La liste est trop grande, donc nous renvoyons le lecteur aux pages 23 à 25 du livre de Beauchamp.

**Perception guidée :** l'ambiguïté des images dans beaucoup de plans filmiques fait que le spectateur cherche dans le son des clés pour construire le sens. Dans ces cas, le design sonore guide la perception du public. Beauchamp invoque le concept de « valeur ajoutée » proposé par Michel Chion.

Bref, ce que nous propose Robin Beauchamp est toute une taxonomie des figures, dans le sens rhétorique, et des rôles du son dans la production audiovisuelle. Et bien qu'il parle surtout pour les réalisateurs –comme les traités classiques de rhétorique parlaient pour des poètes et des orateurs-, et que ses idées s'inscrivent dans une institution concrète (la production industrielle des Studios américains), la description détaillée que nous offre cet auteur nous permettra d'identifier quelques figures (maintenant au sens sémiotique) du son dans le cinéma d'animation. Beauchamp consacre un chapitre entier à chaque élément de la bande-son. Dans les prochains chapitres de notre travail, nous reviendrons à quelques-uns.

### ***Coloriture : une exploration sémiotique et narratologique du son dans l'animation***

Pour finaliser ce parcours de la faible bibliographie consacrée au son dans l'animation, nous voulons évoquer l'œuvre italienne *Coloriture* (BENDAZZI, et al., 1995). Ce travail est très important dans la mesure où c'est presque le seul ouvrage sur le sujet qui adopte une

---

<sup>69</sup> On sent toujours ici l'influence de Michel Chion.

approche plus proche de la sémiotique, au moins dans les chapitres de la première partie : *la théorie*. Mais même dans les deux autres parties il y a toujours des références à la narratologie et autres théories proches des sciences du langage. Pour l'instant nous ne traiterons que de la première partie.

Le premier chapitre de cette partie est *Rumori di Carta : per una teoria del sonoro nell'animazione (bruit de papier: pour une théorie du son dans l'animation)*, de Renato Candia. Dans cet article Candia se propose d'identifier la contribution sémantique du composant sonore au cinéma d'animation, en essayant aussi de formuler une hypothèse sur la spécificité du son dans ce type de cinéma. Dès le début, Candia souligne que cette spécificité a à voir avec la question de la référencialité tant de l'image que du son. Il faut dire que l'auteur parle surtout du « Cartoon », c'est-à-dire du dessin animé de la tradition américaine. Dans ce « contexte expressif », d'après Candia, règne la déformation (bien qu'on puisse dire la même chose de la plupart des techniques de l'animation), et cela est la différence principale par rapport au cinéma PVR<sup>70</sup>.

En chemin, Candia conteste l'idée selon laquelle le rôle du son dans l'audiovisuel est la seule redondance de l'image :

*Si le langage du cinéma est appelé audio-visuel c'est parce qu'il se base, en fait, sur la sensorialité auditive et visuelle, et ce dualisme aide raisonnablement à penser que les éléments de la structure filmique s'organisent selon deux catégories générales: sonore et visuelle. Ces deux catégories portent chacune son propre contenu d'informations que le film, en tant que produit fini, rend homogène et uniforme en raison des propriétés structurales de ce langage. (CANDIA, 1995, pp. 26 – 27)<sup>71</sup>*

Candia cite Chion quand il parle de la fin du cinéma muet en indiquant l'apparition d'un plan structural dans lequel s'établit une syntaxe particulière des relations possibles entre l'image cinématographique et la voix. D'un point de vue narratologique, en suivant André

---

<sup>70</sup> Voir note No. 1.

<sup>71</sup> “*Se il linguaggio del cinema si definisce audio-visivo è perché esso agisce, appunto, sulle sensorialità uditive e visive, e questo dualismo aiuta ragionevolmente a pensare le componenti della struttura filmica organizzate secondo due categorie generali: quella sonora e quella visiva. Queste due categorie sono ciascuna portatrice di un proprio contenuto informativo che il film, come prodotto finito, rende omogeneo e uniforme in virtù delle proprietà strutturali del suo linguaggio*”.

Gaudreault et François Jost<sup>72</sup>, Candia souligne l'existence d'un « double récit » que la narration filmique utilise avec l'objectif d'une désambiguïisation réciproque : ce qu'« on voit » et ce qu'« on entend ».

Dans le cas de l'animation, cette articulation entre le son et l'image est passée par plusieurs étapes différentes. Pour Renato Candia la pratique déjà mentionnée du « *mickey mousing* »<sup>73</sup>, c'est-à-dire, la synchronisation exacte entre des sons et des images n'est que « l'enfance » de la syntaxe audiovisuelle des Cartoons, une structuration « banale » et « élémentaire ». Par contre, Candia décrit un deuxième exemple de structuration audiovisuelle dans le Cartoon : le court métrage *Bugs Bunny Rides Again*<sup>74</sup> (1948) de Fritz Freleng pour le studio Warner, dans lequel le lapin célèbre est dans un vieux bar de l'Ouest, et se jette dans Yosemite Sam, qui dit: « Cette ville est trop petite pour nous deux ». Bugs Bunny sort de cadre (et du bar) et on écoute les sons convulsifs d'un chantier. Le lapin se retourne et dit que, maintenant, la ville est assez grande pour deux. Au dehors du Saloon, on peut voir une grande métropole, pleine de gratte-ciels. D'après l'auteur, ici la désambiguïisation réciproque de l'image et le son est complètement différente que dans le « *mickey-mousing* » :

*L'éloignement rapide de Bunny est couvert par le bruit qu'il fait, alors que le produit visuel de ce bruit est visible uniquement dans les plans suivants. Dans la pratique, la mise en scène a été construite grâce la séparation du bruit de sa source visuelle et cela suffit pour créer une condition d'attente, un désir de la solution de cette séparation, avec la connaissance de la source sonore (dans le film de Freleng la source est, effectivement, la ville des gratte-ciels) (...) l'information, pour ainsi dire, référentielle, oscille de la représentation visuelle à la représentation sonore, pour revenir à la visuelle, et cette oscillation produit une multiplication de l'espace de la représentation, en renvoyant le spectateur à un double ordre d'espaces, ceux qui sont*

---

<sup>72</sup> Cité par Candia. Voir (GAUDREULT & JOST, 1990, pp. 28-30).

<sup>73</sup> Dans cette œuvre il y a un autre article d'Alberto Boschi intitulé « *Éloge du Mickey mousing : l'expression sonore dans les cartoons disneyens* » (BOSCHI, 1995).

<sup>74</sup> Malheureusement ce film n'est pas disponible sur Internet.

*montrés et ceux qui sont évoqués (dans ce cas, par les bruits). (CANDIA, 1995, p. 27)<sup>75</sup>*

Renato Candia affirme que la capacité du son de nous renvoyer au lieu d'une source sonore et d'être signe de cette source devrait être suffisante pour démontrer la non-subordination du son à l'image dans le cinéma, ainsi que les capacités dramatiques et expressives de la relation audiovisuelle, que l'auteur met en relation avec l'idée de Noël Burch d'un « simulacre total » : la représentation pleine d'un monde à travers des sons et des images. Sur ce point, Candia se demande si dans l'animation ce pouvoir de simulation opère de la même façon. En principe, la réponse à cette question est affirmative, puisque le cinéma d'animation a été constitué au cœur du canon du langage audiovisuel.

Néanmoins, Candia croit que la tendance à la déformation dans l'animation est une caractéristique distinctive. En ce sens, tandis que le cinéma PVR se rapproche du récit littéraire, l'animation se rapproche de la fable. Et ici nous arrivons à une des idées les plus intéressantes de cet article : d'après Candia, dans le cinéma d'animation la représentation visuelle de la réalité passe à travers un double niveau de référencialité : ce qui renvoie le dessin à la réalité, et ce qui renvoie le film au dessin ; d'où le rapport avec la fable : « l'animation active un régime narratif qui est étrange par sa nature même, où tout est déjà possible dès le début et au-delà des limites de la physique des corps et des matériaux de la réalité quotidienne » (CANDIA, 1995, p. 30)<sup>76</sup>. Mais pour le son les choses sont différentes :

---

<sup>75</sup> *“Il rapidissimo allontanamento di Bunny è coperto dal rumore di ciò che fa, mentre il prodotto visivo di quello stesso rumore è mostrato soltanto nell'inquadratura successiva. In pratica la messa in scena è stata costruita attraverso una separazione del rumore dalla sua fonte visiva e questo è sufficiente a creare una condizione d'attesa, un desiderio verso la soluzione di questa separazione, cioè verso la conoscenza della fonte sonora (che nel film di Freleng è appunto la città fatta di grattacieli) (...) L'informazione, per così dire, referenziale oscilla dalla rappresentazione visiva a quella sonora per ritornare a quella visiva, e questa oscillazione crea una evidente moltiplicazione degli spazi della rappresentazione rinviando lo spettatore ad un doppio ordine di luoghi, quelli mostrati e quelli evocati (in questo caso dal rumore)”.*

<sup>76</sup> *“L'animazione attiva un regime narrativo che è strano per sua stessa natura, dove tutto è già possibile fin dall'inizio e ben oltre i limiti della fisicità dei corpi e dei materiali della realtà quotidiana”.*

car la double référencialité du visible<sup>77</sup> dans le cinéma d'animation est condensée dans un seul rapport de référencialité sonore dont les termes sont le film (qui représente) et la réalité (qui est représentée). La manipulation de cette asymétrie donne lieu à tout type de possibilités symboliques et narratives.

Pour illustrer cette caractéristique Renato Candia pose un exemple très intéressant, cette fois de l'animation d'objets (poupées)<sup>78</sup> : le court-métrage *Jasper in a Jam* (1946) du réalisateur George Pal<sup>79</sup>. Dans cette production, le petit Jasper entre chez un prêteur sur gages pour se protéger de la pluie. Là, quand il tente d'échapper, la porte se ferme, les objets prennent vie et commencent à jouer un jam de jazz. Candia souligne qu'est un effet sonore ce qui marque la transformation merveilleuse des objets : le son d'un coucou à minuit. Jusqu'ici, tous les sons sont réalistes. Par contre, le son de l'horloge est complètement exagéré : en effet, ce n'est pas le son d'une horloge, mais d'un mécanisme industriel, d'un atelier entier de réparations. À partir de ce point, la sonorité change et la musique s'impose, une musique qui se joue elle-même, autant que les instruments sont animés. Voici l'interprétation de Candia :

*La soudaine folie visuelle et sonore de l'horloge-coucou indique, le cas échéant, la variation du registre narratif, qui passe d'une fausse « réalité » à la fable proprement dite. (...) Ces exemples semblent conduire à la considération d'un niveau de référencialité exclusif du son qui, en sautant complètement le niveau intermédiaire, vient établir dans le film une représentation directe de la réalité : il accomplit cela dans l'espace-temps qu'il occupe dans*

---

<sup>77</sup> Une idée partagée par Iuri Lotman : « une propriété au début du langage du cinéma d'animation c'est qu'il fonctionne avec des signes de signes : ce qui se passe face au spectateur dans l'écran c'est une représentation d'une représentation. En plus, si le mouvement double la capacité de produire l'illusion de la photographie, la photographie elle-même double la conventionnalité de l'image dessinée » (LOTMAN, 2000, pp. 139-140).

<sup>78</sup> Ce type d'animation est différent du dessin animé, car il manque de la double référencialité dont nous parlons. Néanmoins, l'auteur se demande si : « est-ce que sa matérialité concrète, déjà forgée par la reprise directe, guide (sinon détermine de manière exhaustive) la référencialité du texte sonore sur lequel elle opère ? » (« *è la loro concretezza materiale, pur già contraffatta dalla ripresa a passo uno, a orientare (se non proprio a qualificare esaurientemente) la referenzialità del testo sonoro nel quale vengono fatti agire?* »). p. 30.

<sup>79</sup> Pour audio-visioner le film : <https://www.youtube.com/watch?v=TYF7Zx8G51g>

*l'histoire du film, en réduisant le bruit de la réalité à la recherche de ses propres règles spécifiques de représentation, de sa propre capacité expressive. (CANDIA, 1995, p. 31)<sup>80</sup>*

L'auteur affirme que ce pouvoir dérivé de la manipulation du son a aussi des conséquences très importantes dans la dimension de l'énonciation, en permettant des effets inédits au-delà de l'image.

Dans la dernière partie de son texte, Renato Candia développe tout une typologie des usages du son dans le cinéma d'animation, d'après le graphique proposé par Michel Chion dans *Le son au cinéma*, pour montrer la manière particulière dans laquelle l'animation redessine les frontières entre ces catégories : son in/hors-champ/off, zones acousmatiques et zone visuelle. Cette fois l'exemple est un autre classique du cartoon : la série du Coyote et du Beep beep. L'auteur illustre comment dans plusieurs épisodes le dessin animé se déplace parmi ces frontières, en donnant lieu à des jeux inédits. À différence de la plupart des textes anglo-saxons, cette petite étude permet d'identifier plusieurs noyaux théoriques importants pour l'étude du son dans l'animation d'un point de vue proprement sémiotique.

### **Maria Luisa Bionda : la narrativité de la musique dans le cartoon**

Le deuxième article du livre offre aussi quelques idées en rapport avec la sémiotique : le texte de Maria Luisa Bionda nous propose quelques « notes pour une approche narratologique » du son dans le Cartoons (BIONDA, 1995). Cet article soutient tout un dialogue avec le texte de Candia, qui lui sert de fondement. Elle évoque aussi la mal comprise « redondance » du son en rapport à l'image, en affirmant qu'elle ne doit pas être conçue toujours comme une limite, mais comme une potentialité, à cause de sa dimension évocatrice. Elle parle aussi du « simulacre total » du cinéma, mais note que dans le cas de l'animation il s'agit plus d'un « simulacre virtuel », dans le sens que le dessin animé construit des personnages et des événements impossibles dans un univers qui soit une simple représentation

---

<sup>80</sup> *“L'improvvisa follia visiva e sonora dell'orologio a cucù segnala, semmai, la variazione del registro narrativo che passa da un finto “dal vero” alla fiaba vera e propria. (...) Questi esempi sembrano portarci alla considerazione di un único livello di referenzialità del sonoro che, saltando completamente ogni livello intermedio, giunge a stabilire dal film un rapporto diretto di rappresentazione del reale: compie cioè nello spazio-tempo che esso occupa nella vicenda del film, una sua riduzione dei rumori della realtà alla ricerca di una sua specifica normativa di rappresentazione, di una sua personale capacità spressiva”.*

du réel, mais qui peuvent donner lieu à « un univers potentiel d’histoire infinie », plus qu’à une simple référentialité diégétique. Et le son y joue un rôle fondamental :

*Le son peut être anaphorique par rapport aux événements qu’il accompagne. Le son peut anticiper, convaincre, évoquer. Peut multiplier ou diminuer la représentabilité de la réalité et de l’irréalité en faisant écouter des sons d’un événement qui ne se déroule pas en premier plan, mais dont on peut pressentir qu’il aura son importance dans le développement de l’histoire. La narration peut osciller de la dimension sonore à la dimension visuelle en amplifiant ultérieurement la portion d’univers, réel ou irréel, qu’elle soit représentée ou représentable. Le son rend représentable l’imaginaire et en même temps, rend imaginaire ce qui est en train d’être représenté. (BIONDA, 1995, p. 46)<sup>81</sup>*

Encore une fois, Bionda utilise la comparaison avec le modèle « Disney » pour montrer d’autres possibilités expressives et narratives du son dans le Cartoon. Ainsi, elle décrit l’usage de la musique et du son dans toute la production des studios Disney, de *Fantasia* –qui, à son avis, n’est pas un prodige de musicalisation, mais l’imposition de la simple synchronisation comme la seule possible relation entre les images et la musique– jusqu’à *La belle et la bête* (1991), un des derniers films lancés par le studio à son époque. Ici, le son a la seule fonction de redondance par rapport à l’image, afin de créer un univers fantastique, mais vraisemblable. Bionda souligne, par exemple, les sons des objets animés du film :

*bien que, dans ce cas, pour n’avoir rencontré dans l’univers réel aucune correspondance référentielle, les sons des objets animés sont en fait imprégnés de réalisme (...) tous les objets animés deviennent humains lorsque le sortilège est rompu, mais il ne s’agit pas d’une transformation, comme dans le cas de la bête, il s’agit d’un simple retour à l’origine, où les caractéristiques essentielles des personnages est leur humanité, et n’ont été jamais perdues, seulement cachées, ils sont transportés de retour à la normalité, en conservant leur voix, qui reste identique lorsque le sortilège est actif, et lorsqu’il est rompu. Tout reste normal, en*

---

<sup>81</sup> “Il suono può essere anaforico rispetto agli avvenimenti che seguiranno. Può anticipare, convincere, evocare. Può moltiplicare a dismisura la rappresentabilità del reale e dell’irreale facendo udire suoni di avvenimenti che non accadono in primo piano, ma che si intuisce avranno un peso fondamentale nello sviluppo della storia. La narrazione può oscillare dalla dimensione sonora alla dimensione visiva amplificando ulteriormente la proiezione di universo, reale o irreale che sia, rappresentata e rappresentabile. Il sonoro rende rappresentabile l’immaginario e contemporaneamente rende immaginario ciò che viene rappresentato”.



*somme, parce que ce qu'on cherche à obtenir avec l'usage des bruits et de la voix off c'est un appui de la réalité, ou mieux, de l'hyperréalité, en excluant l'indépendance des sons et des images. (BIONDA, 1995, pp. 48-49)<sup>82</sup>*

À cette utilisation redondante et intégrale des sons et des images s'oppose l'univers animé de Tex Avery. Bionda identifie aussi des oppositions sémantiques : si l'univers de Disney est optimiste et hanté, l'univers d'Avery est pessimiste et ironique. Et ces différences ont leur correspondance dans le son :

*La production d'un son non réaliste par un objet transformé pour qu'il ne soit pas réaliste permet au narrateur de créer de nouvelles catégories, un langage nouveau qui montre la réalité comme on la présente et la sent en interne, plus que comme on la voit. C'est comme si les sons et les bruits absorbaient la fonction de comprendre que la vie n'est pas seulement concrète, mais qu'elle est aussi rêve et désillusion, folie et calme, résignation et révolte. (BIONDA, 1995, p. 50)<sup>83</sup>*

M. L. Bionda consacre la dernière partie de son article à souligner quelques différences entre la fonction diégétique de la musique dans deux types de séries télévisées. Les émissions à récits complets, qui sont résolus dans un seul épisode, ont un mode musical constant qui représente les humeurs des personnages d'une façon homogène. Par contre, dans les séries à épisodes ouverts (plus proches des téléromans), la musique est modulée d'après les changements des états d'âme des personnages, dans une sorte « d'annihilation de la

---

<sup>82</sup> *“anche in questo caso, pur nono trovando nell'universo reale alcuna corrisponendenza referenziale, i rumori degli oggetti animati sono di fatto impregnati di realismo. (...) tutti gli oggetti animati ritornano persone quando l'incantesimo si rompe, ma non si tratta di una conversione, come nel caso della Bestia, si tratta di un semplice ritorno alle origini dove sia le caratteristiche essenziali dei personaggi sia la loro umanità, mai perdute, ma solo celate, sono traghettate verso il ritorno alla normalità con la conservazione della voce ch rimane identica sia durante l'icantesimo sia quando l'incantesimo viene spezzato. Tutto nella normalità, insomma, perché quello che si vuole ottenere con l'uso dei rumori e delle voci off è un ausilio di realtà, o meglio di iper-realtà, escludendo la possibilità di uno scollamento di sonoro e immagini”.*

<sup>83</sup> *“La produzione di un suono non realistico da parte di un oggetto trasformato in modo altrettanto non realistico permette a chi racconta di creare delle nuove categorie, un nuovo linguaggio, che mostri la realtà come la intuiamo e come la sentiamo interiormente, piuttosto che come la vediamo. È come se suoni e rumori assolvessero la funzione di far intendere che la vita è altro, non è solo concretezza, ma è anche sogno e disincanto, follia e pacatezza, rassegnazione e ribellione”.*



dimension extérieure ». Ce type de sonorité démontre l'autonomie et l'importance du son dans la narration audiovisuelle :

*Le spectateur se sent maître de la représentation parce qu'en l'attendant, il peut la dominer. Nous sommes donc confrontés à une utilisation quasi syntaxique de la dimension sonore, une utilisation qui détermine le récit, qui marque ses caractéristiques particulières. (BIONDA, 1995, p. 52)<sup>84</sup>*

### **Rita De Giuli et son compte-rendu de Jon Newsom**

Le texte de Rita De Giuli est une sorte de compte-rendu du texte de Jon Newsom, *A Sound Idea: Music for Animated Films* (NEWSOM, 1980) lequel est, pour l'auteure, « un excellent précurseur des analyses sur les relations son-image dans le secteur du cartoon américain » (DE GIULI, 1995)<sup>85</sup>. De Giuli s'arrête sur la richesse de l'analyse que Newsom fait de la séquence « *Flight to Neverland* » du film *Peter Pan* (Walt Disney, 1953), dont l'auteur montre comment la complexité n'est pas due seulement à l'utilisation de la caméra multiplane pour la construction de l'image en profondeur, mais aussi à la bande musicale, qui a nécessité la réalisation de nombreux segments séparés puis combinés dans le doublage final. La combinaison des paroles, les effets sonores et de la piste musicale avec quatre niveaux de profondeur de l'image dans la camera multiplane en parfait synchronisme est le grand triomphe de cette séquence.

Newsom offre aussi une petite histoire des expérimentations audiovisuelles dans l'histoire du cinéma, comme le *Ballet Mécanique de Fernand Léger*, dont la musique de George Antheil, a été composée pour être jouée en synchronisme avec le film ; ou le film *Joie de vivre* (1934) d'Hector Hoppin et Anthony Gross, dont nous reparlerons dans la perspective de Maureen Furniss.

Ensuite, Newsom consacre une grande partie de son texte au cas de Disney comme pionnier du son et de la musique dans le cinéma d'animation. En plus du film *Steamboat*

---

<sup>84</sup> “*Lo spettatore si sente padrone della rappresentazione perché prevedendola la può dominare. Ci troviamo quindi di fronte ad un uso quasi sintattico della dimensione sonora, un uso che determina la narrazione, che ne segna le caratteristiche peculiari*”.

<sup>85</sup> “*Gli eccellenti precursori dell'analisi del rapporto suono-immagine nel partitorari settore del cartoon americano*”.

Willie, Newsom mentionne la *Silly Symphony Three little pigs*, pour souligner comment la musique de Frank Churchill sert à construire et développer l'histoire<sup>86</sup>. Newsom raconte aussi le développement de la technologie Vitaphone, système choisi par Disney, et sa relation avec le Phonofilm de Lee De Forest. En plus, il fait un parcours détaillé pour l'histoire de l'usage de la musique dans les différents Studios d'animation américains, comme MGM, où Scott Bradley développa sa *funny music*, ou l'UPA, l'espace de travail de David Raskin. Bref, le texte de Newsom est un des premiers travaux rigoureux sur la musique et le son dans les cartoons, en anticipant des œuvres plus récentes comme celles de Goldmark. De Giuli conclut son compte rendu en soulignant l'importance historique du texte de Newsom :

*Son mérite c'est d'avoir considéré l'histoire du cinéma d'animation comme l'histoire d'un art musical, précisément pour avoir trouvé sa vraie dimension avec l'invention du cinéma sonore. La complexité des rapports entre la musique et le film d'animation est explorée dans le texte, comme on a vu, non seulement dans sa phase d'élaboration en relation avec l'histoire et les personnages, mais par rapport à la synchronisation avec l'image, dans le rôle fondamental de la création des atmosphères, sentiments et émotions, qui sont le véhicule authentique de nombreuses histoires. (DE GIULI, 1995, p. 43)*

## **Giuseppe Valenzise et le pluricode de l'animation**

Dans cet article, Valenzi fait une description détaillée des composants expressifs du cinéma d'animation, entendus comme des « codes » provenant d'autres langages : le langage cinématographique avec des cadrages, des mouvements de la caméra, le montage, etc. ; le langage graphique du dessin (et, particulièrement, de la caricature) ; le propre langage de l'animation, entendu comme la représentation dynamique et la mise en mouvement du dessin ; et les langages de la musique et des paroles. Quant à la musique, Valenzi remarque :

*Dans un dessin animé traditionnel, la musique a des fonctions diverses et plus importantes que dans un film de « PVR ». Dans ce dernier elle sert d'accompagnement et d'exaltation des moments dramatiques particuliers ou à créer l'atmosphère d'une scène. Dans le cartoon la*

---

<sup>86</sup> Dans son livre *Film Rhythm after Sound. Thecnology, Music and Performance*, Lea Jacobs analyse en détail le rôle de la musique dans les *Silly Sinphonies*, et s'arrête surtout au cas du film *Three little pigs*. L'auteur fait une description pas à pas de la partition des chansons (notamment, le thème *Who's affraid of the Big Bad Wolf ?*, écrit par Frank Churchill) par rapport aux mouvements des personnages animés, et du rythme visuel du film. Jacobs analyse aussi *Playful Pluto* (Disney, 1934). Voir JACOBS (2015, pp. 85-89).

*musique trouve une fonction similaire à ce qu'elle occupe dans le ballet classique. C'est sur la base de la musique, en effet, que l'animateur fait marcher, courir et mobiliser ses personnages.* (VALENZISE, 1995, p. 54)

Valenzise explique la technique de synchronisation, et le travail des musiciens qui ont composé ou adapté des musiques pour le cartoon : Carl W. Stalling, Frank Churchill, Paul Smith, Oliver Wallace, George Bruns, Sammy Timberg, Walter Lantz où Scott Bradley, en essayant de montrer l'évolution de la musique, des adaptations de pièces classiques à l'introduction des genres du jazz plus contemporain. Un paragraphe intéressant de ce texte est celui où Valenzise offre une taxonomie des différentes façons dont le cartoon prend la musique comme sujet : 1) le discours filmique présente une abstraction animée sur la base de la musique (*Fantasia*, 1940) ; 2) le discours filmique présente une narration floue, création d'une atmosphère visuelle à travers des images suggestives, sans narrer une histoire précise (*Skeleton Dance*, 1929, à partir de la *Dance Macabre* de Saint-Saens). 3) le discours filmique narre une histoire précise adaptée à la composition musicale (la scène de *L'Apprenti Sorcier* dans *Fantasia*) ; 4) le discours filmique narre l'histoire d'un chanteur, un directeur d'orchestre, un compositeur ou musicien de musique populaire, classique ou du jazz (*The jazz concert*, de Disney (1935) ou *Rhapsody Rabbit* (1946) de Warner bros.) (VALENZISE, 1995, pp. 59-60).

Dans le paragraphe consacré aux voix dans le cartoon, Valenzise nous rappelle le fait que même les comédiens du cinéma muet (Chaplin, Keaton) n'ont pas immédiatement donné leur voix à leurs personnages. Les personnages du dessin animé n'ont pas non plus parlé immédiatement, et la question du « lip sync » a été un des premiers problèmes. Néanmoins, une fois réussie, la caractérisation des personnages à partir de la voix est devenue la pratique la plus étendue. Valenzise raconte un commentaire de Tex Avery sur l'importance de la voix pour le succès des personnages animés comme Donald Duck, Popeye ou Droopy. Valenzise aussi consacre quelques lignes à l'usage « métalectique » de la voix-off dans le cartoon, comme dans le film *Red Hot Riding Hood* de Tex Avery, où un personnage dialogue avec la voix-off du narrateur. L'auteur décrit aussi des problèmes dérivés du doublage.

Ensuite, Valenzise s'arrête sur la question des « effets sonores » dans le cartoon, en se demandant s'il est possible de parler d'un son « déformé » ou « caricaturé ». Il cite Bella Balazs, pour qui ce n'est pas le cas. D'après Balazs, la déformation visuelle est possible parce qu'on se souvient de la « forme originelle » qui est déformée, tandis que dans le son on ne

perçoit qu'un « mode », une manifestation. Il est toujours possible à jouer avec la (non) correspondance comique entre un son et une source, si on donne le rugissement d'un lion à une souris, et cela produit un effet comique. Mais il n'y a pas de distorsion.

Par contre, Valenzise croit que cela est toujours possible, surtout dans le cas de sons très stéréotypés, si l'on les ralentit ou les accélère, même en absence de la source sonore. Dans ce cas, l'auditeur peut reconnaître la déformation et cela peut produire un effet comique du son lui-même. Valenzise raconte des anecdotes des bruiteurs traditionnels et nous rappelle la richesse du « foley » dans l'animation. Il mentionne aussi le recours très commun dans le cartoon, à la non-correspondance entre sons et sources visuelles qui donne au cartoon sa sonorité particulière. À ce sujet, il mentionne un cas très intéressant, où le protagoniste d'un court métrage de Tex Avery, un turbulent écureuil, devient bruiteur et produit à l'écran les effets sonores de l'accident de son antagoniste et victime, un chien malheureux, à travers des tambours, grosses caisses, cymbales, sifflets, crécelles et plus encore : il s'agit de *Screwball Squirrel* (Tex Avery, 1944)<sup>87</sup>.

Valenzise finit son article avec une réflexion intéressante sur la difficulté de différencier les voix, les musiques et les bruits dans les cartoons : « chaque élément se fond créativement dans une tessiture acoustique dont les composants sont inclassables selon la distinction musique, voix, effets sonores » (VALENZISE, 1995, p. 69). L'auteur fait également un rapprochement entre la sonorité du cartoon et certaines expériences de la musique concrète, comme l'*Étude aux casseroles*, de Schaeffer (1948). Finalement, Valenzise affirme que, si l'on sort du territoire du cartoon sériel, pour explorer l'animation indépendante, il sera plus encore évident que la fusion audio-visuelle a conquis ici « le plus haut degré de sophistication et de l'efficacité artistique » (Ibid., 71).

### **Alberto Boschi et l'éloge du *Mickey-mousing***

Dans le dernier chapitre de *Coloriture* qui m'intéresse, Alberto Boschi propose une comparaison entre les stratégies sonores des dessins animés et celles du cinéma PVR, en soulignant l'influence des premiers sur le second. Fondé surtout sur les idées de Michel Chion, Boschi revient sur la pratique de bruitage dans les cartoons des années 30 et 40. D'une façon similaire à celle de Valenzise, Boschi compare la construction sonore des cartoons avec

---

<sup>87</sup> Pour audio-visioner ce film: <https://vimeo.com/145808279>

une « partition semi-concrète », dans la mesure où on trouve des compositions musicales avec insertion de bruits, pour faire des accents où des surprises. D'après Boschi, si dans *Steamboat Willie* la conception sonore n'avait pas atteint sa maturité, on peut la trouver dans le travail d'Ub Iwerks hors du studio Disney. Particulièrement, Boschi parle du film *The Cukoo Murder Case* :

*Tandis que la bande-son de Steamboat Willie produit à l'écoute l'impression constante de que la musique et les bruits sont disposés à deux niveaux différents, dans The Cukoo Murder Case le bruitage se révèle plus solidaire avec la partition (composée par Carl Stalling) et il semble être placé sur le même plan sonore. Si, en effet, la musique -synchronisée beaucoup mieux avec l'action visuelle et, par conséquent, sans cesse soumise à des accélérations, décélérations, des arrêts brusques et des changements de rythme- est interrompue brièvement pour laisser libre l'espace pour les effets sonores, toutefois, ces effets -beaucoup plus soumis au rythme imposé par la musique- viennent s'incorporer à l'accompagnement de l'orchestre.* (BOSCHI, 1995, p. 77)

Boschi nous rappelle qu'une bonne partie des sons diégétiques dans le cartoon classique viennent de la musique, dans une pratique partagée avec le cinéma PVR des premiers temps. Boschi revient enfin sur le *mickey-mousing* et l'*underscoring*. L'auteur cite Curtis (dont on a déjà parlé) et son idée de l'isomorphisme entre la bande-son et la bande-image, pour revendiquer la pratique du *mickey-mousing*. À son avis, la critique de cette pratique vient de la relégation de la musique au deuxième plan dans le cinéma PVR, que le cartoon des années 30 et 40 bouleverse, en redonnant au cinéma la force rythmique de la musique et la richesse sonore des bruits. Ainsi, le cartoon recouvre le pouvoir expressif du muet, en offrant une alternative au verbocentrisme du cinéma parlant.

À mon avis, les articles recueillis dans *Coloriture* sont un échantillon des possibilités qu'offre l'animation pour une étude théorique sérieuse. Je ne me suis arrêté que sur la première partie du livre à cause de son caractère plus théorique. Néanmoins les analyses pratiques que cet ouvrage propose sont aussi importantes. Par exemple, en plus de l'accent marqué sur le cartoon hollywoodien dans cette partie (et d'autres articles aussi), le livre inclut des articles sur la relation entre le dessin animé et la musique classique et populaire (notamment, le pop-rock et le jazz) ; sur des animateurs indépendants comme Len Lye, Norman McLaren, Norman Royer ou Raymond Scott ; sur des réalisateurs italiens d'animation comme Osvaldo Cavandoli, Emmanuele Luzzati et Giulio Gianini ; sur les effets et la musique dans l'animation japonaise de genre cyberpunk ; et sur la synesthésie dans le

cinéma d'animation. En plus, *Coloritore* est l'ouvrage où la sémiotique a une présence plus forte, avec une inclination vers la narratologie qui donne à ce recueil son unité théorique.

### **Daniel Goldmark et la musique des cartoons**

Dans une ligne similaire se trouve l'œuvre de Daniel Goldmark, un musicologue qui a consacré son travail à l'étude de la musique et à la sonorité des cartoons américains. Son livre le plus vaste à ce sujet est *Tunes for 'Toons. Music and the Hollywood Cartoon* (GOLDMARK, 2005), où l'auteur propose un ensemble d'analyses de cas sur des musiciens et tendances musicales dans le cartoon américain classique, y compris des chapitres consacrés à deux des compositeurs de musique d'animation les plus célèbres : Carl Stalling et Scott Bradley ; un chapitre sur la relation entre l'animation, le jazz et la culture du swing ; un autre chapitre sur la musique classique et les cartoons et une analyse du court métrage *What's Opera, Doc ?* (Chuck Jones, 1957)<sup>88</sup>.

Tandis que les autres chapitres sont plus concentrés dans des aspects musicaux où biographiques, ce dernier offre une analyse de toutes les dimensions du court métrage : l'histoire, la présence de la musique de Wagner dans le cartoon en général et dans ce film en particulier, le contexte historique de la présence de l'opéra dans le cartoon, la composante visuelle du film, la musique –bien sûr-, et la *narration* (storytelling), ce qui fait de cette étude une véritable lecture narratologique qui cherche à démontrer la versatilité narrative des réalisateurs, qui réussissent à intégrer le comique des cartoons avec le caractère dramatique de Wagner, comme on peut voir dans ce passage :

*Bien que la musique omniprésente de Tannhäuser soit importante dans What's Opera, Doc ? la mélodie la plus mémorable et captivante du court métrage c'est le leitmotiv de la Valkyrie, adapté et fusionné avec le cri de bataille d'Elmer : « Kill the Wabbit ! ». Elmer fait sa déclaration dans une des scènes plus courtes, mais cette phrase a duré de façon plus indélébile que tout autre dans la mémoire collective d'innombrables téléspectateurs de dessins animés. L'expression « Kill the wabbit » encapsule presque deux décennies de lutte pour Elmer (il a commencé la chasse de Bugs en 1940), et elle est greffée sur un leitmotiv qui illustre parfaitement la forme: courte, distincte, unique, et donc facile à identifier. Intégrer la quintessence de Wagner avec l'essence d'Elmer était une idée brillante ; cela est si efficace, en fait, que ceux qui voient le cartoon semblent incapables de déchirer les deux icônes, l'une*

---

<sup>88</sup> Pour audio-visionner le film : <https://vimeo.com/115773980>

visuelle, l'autre sonore. Ce geste musical signifie encore Wagner, mais il est maintenant Wagner chez Bugs et Elmer. (GOLDMARK, 2005, p. 152)<sup>89</sup>

L'intérêt de Goldmark dans la musique des cartoons continue dans le recueil qu'il a édité avec Yuval Taylor, un ensemble d'interviews, essais et fragments de textes d'animateurs, compositeurs, musiciens, et personnages du monde de la musique et des cartoons (GOLDMARK & TAYLOR, 2002). Ici Goldmark fait des observations au sujet du livre et des interviews avec des personnages comme Mark Mothersbaugh, Alf Clausen, Richard Stone, Steve Bernstein et Julie Bernstein, dans une véritable ethnographie de la musique « animée ».

Goldmark a édité aussi le livre *Funny Pictures: Animation and Comedy in Studio-Era Hollywood* avec Charles Keil (GOLDMARK & KEIL, 2011). Dans ce recueil, Goldmark participe avec l'étude intitulée *Sounds Funny / Funny sounds : Theorizing Cartoon Music*, où il revient sur le *mickey-mousing*, pour montrer son retour dans les dessins animés des dernières années. Dans ce texte Goldmark aussi offre un commentaire sur la visualisation de la musique dans les cartoons classiques, avec plusieurs exemples dont le plus intéressant est celui du film *The Herring Murder Mystery* (Dun Roman, 1943)<sup>90</sup>, dans lequel il y a un moment où le son ne s'écoute pas, et un personnage (le juge) demande : « est-ce qu'il y a un problème avec la bande-son ? (*"Is anything wrong with the soundtrack?"*) ». Alors sur l'écran apparaît le bord de la pellicule avec la bande-son, et l'onde graphique prend forme humaine pour répondre : « non, il n'y a pas de problème avec la bande-son, il suffit de parler un peu

---

<sup>89</sup> "Though the ubiquitous music from *Tannhäuser* is prominent in *What's Opera, Doc?*, no melody in this short is more memorable and arresting than the *Valkyrie* leitmotif, appropriated and then fused with Elmer's hunting cry of "Kill the wabbit!". Elmer makes this declaration in the briefest of scenes, yet this phrase has lasted more indelibly than any other in the collective memory of countless cartoon viewers. The phrase "Kill the wabbit" encapsulates almost two decades of struggle for Elmer (he began hunting Bugs in 1940), and it is grafted onto a leitmotif that perfectly exemplifies the form: short, distinct, unique, and thus easy to identify. Uniting the epitome of Wagner with the essence of Elmer was a brilliant idea; it is so effective, in fact, that those who see the cartoon seem unable to tear apart the two icons—one visual, one aural. This musical gesture still signifies Wagner, but it is now Wagner as portrayed by Bugs and Elmer".

<sup>90</sup> Dans L'Internet Movie Data Base (IMDB), le film est daté en 1944. Pour audio-visionner ce film : <https://www.youtube.com/watch?v=q3uOapx20hY>



plus distinctement, stupide ! (“*No, nothing’s wrong with the soundtrack. Just speak a little more distinctly, stupid!*”) » (Voir figure 6).

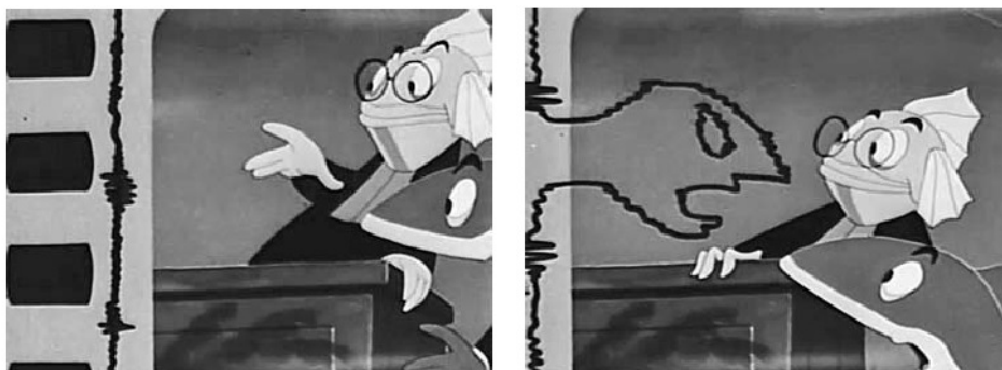


Figure 6 : Photogrammes du film *The Herring Murder Mystery* (1943)

Source : [GOLDMARK, 2011]

Goldmark remarque que « la stratégie qui consiste à représenter visuellement la bande sonore –sans mentionner le fait de converser avec elle- signifie que les créateurs du dessin animé reconnaissent une partie normalement invisible de la mécanique de l’animation », en étendant la tradition de se moquer des processus de production des cartoons eux-mêmes. Il s’agit d’une forme de ce que Gérard Genette appelle *métalepse*, dont on va parler dans le chapitre 4.

Goldmark mentionne le livre de Claudia Gorbman, *Unheard Melodies* (GORBMAN, 1987), où l’auteure défend l’idée que dans les films narratifs la musique doit rester « invisible », de sorte qu’elle ne fasse pas obstacle à l’attention qu’on devrait diriger sur l’histoire. Après le film mentionné Goldmark propose d’ajouter à la liste des caractéristiques qu’on peut attendre de la musique (et d’usage) dans le cinéma hollywoodien, celle de devenir la cible des moqueries.

Nous reviendrons au chapitre 3 sur un autre article de Goldmark sur la musique dans les cartoons avant l’apparition du son synchronisé (GOLDMARK, 2007), et nous avons déjà parlé de l’article sur *Les triplettes de Belleville*. Tous ces textes montrent l’acuité de cet auteur à l’heure d’identifier des usages et stratégies de production sonore dans les dessins animés, bien qu’il se concentre surtout sur la musique, en décrivant tout son potentiel narratif et expressif.



Avant de passer aux conclusions de ce chapitre, je voudrais mentionner d'autres textes sur ce sujet de recherche. Il s'agit surtout d'analyses concrètes de films, mais qui donnent aussi quelques clés. Dans l'*Oxford Handbook of New Audiovisual Aesthetics* (RICHARDSON, et al., 2015), il y a toute une section consacrée aux *sons animés* (Animated Sounds), avec un article de Daniel Goldmark sur la bande-son des films du Studio Pixar ; un texte de Randy Thom sur la conception sonore dans le cinéma d'animation contemporain ; et un article de Lisa Perrott sur « la musique visuelle immersive » de Len Lye. Goldmark participe aussi à l'ouvrage *Oxford Handbook of Film Music Studies* (NEUMEYER, 2014), avec un article intitulé *Drawing a New Narrative for Cartoon Music*, dans lequel il fait un parcours historique sur les tendances de la musique dans le dessin animé américain. Finalement, dans le livre *Sounds of the Future. Essays on Music in Science Fiction Film* (BARTKOWIAK, 2009), on trouve l'article *Hello WALL-E!: Nostalgia, Utopia, and the Science Fiction Musical*, de Kathryn A.T. Edney et Kit Hughes, dans lequel les auteurs analysent le film de Pixar Wall-E (Andrew Stanton, 2008), en montrant les relations entre le cinéma de science-fiction et le film musical (EDNEY & HUGHES, 2009).

## Deuxième synthèse théorique

Une première remarque qu'on peut faire tient au fait qu'il y a un groupe majoritaire de recherches ou de réflexions sur le *cartoon classique hollywoodien*. Dans ce sens, les « *animation studies* » répliquent la tendance des « grandes théories du cinéma » ç se concentrer sur le *canon*, c'est-à-dire, sur les expressions les plus populaires et stabilisées dans une culture déterminée. En tout cas, le caractère périphérique de l'animation dans l'intérêt des théoriciens fait que les remarques qu'on peut recueillir sur ce type de cinéma offrent toujours des informations inédites sur les possibilités de combinaisons audiovisuelles. Les études sur les premiers cartoons sonorisés (et même les recherches sur les cartoons muets) montrent l'apparition d'un ensemble de conventions expressives qu'on a même extrait de l'univers de l'animation, au profit des productions de PVR.

Dans les différentes théories et analyses ici recueillies, on trouve quelques constantes qui mettent en évidence des aspects particuliers de l'animation. Une de ces constantes est la suprématie de la musique dans la conception générale des dessins animés, au point qu'une théoricienne comme Furnis en fait une partie intégrale de la *conception structurale*.

Une autre constante est l'identification du caractère flou des frontières entre les types de *matérialité sonore* dans le cartoon. Il n'est pas toujours facile de marquer la limite entre les bruits et les musiques, ou même entre la voix et les bruits. Il semble y avoir un *continuum sonore* qui se décline vers l'un des trois « pôles » traditionnellement identifiés comme constitutifs des « codes sonores » du cinéma.

Une troisième constante théorique est l'idée de l'existence d'un certain *isomorphisme* entre les distorsions ou altérations du flux sonore et les exagérations et déformations visuelles des personnages et objets animés. On peut parler d'une certaine « plasticité » des formes tant visuelles que sonores, une sorte de solidarité dans la déformation.

Une autre idée qui revient souvent concerne la narrativité dans le dessin animé. Il semble que dans plusieurs cas, la structure générale d'un cartoon n'est pas régie par le flux narratif (même dans des productions clairement narratives), mais par des rythmes et dynamiques audiovisuels.

Finalement, plusieurs auteurs ont souligné la tendance à introduire des ruptures dans les frontières de la diégèse, manifestée sous la figure de la *métalepse* dans le sens narratologique du terme. Cette vocation métaleptique de l'animation est portée jusqu'à des limites vraiment inédites, en transgressant même les frontières du support matériel, comme l'illustre très bien l'exemple du film *The Herring Murder Mystery*.

Ainsi, dans le cadre des dimensions proposées à la fin du dernier chapitre, on peut dire que dans le cas de l'animation les dimensions de la technologie, la matérialité et la construction discursive entrent dans des relations inédites dans ce type de films, le son et ses différentes manifestations étant vraiment au centre de l'expérience audiovisuelle. En ce sens, le cinéma d'animation amène au premier plan la question de l'énonciation « vivante », où l'approche de la textualité des films ne peut plus s'éprouver qu'avec le corps.

Maintenant, mon défi est de rassembler toutes ces intuitions, recueillies chez différents auteurs, et selon plusieurs approches, dans un cadre théorique proprement sémiotique. Dans le chapitre prochain, je me consacre à la construction d'un tel cadre, pour proposer ensuite des analyses concrètes de films intentionnellement choisis hors de la tradition du cartoon.

## Chapitre 3. L'énonciation, le corps et les dispositifs audiovisuels

Le cinéma et les médias audiovisuels participent ensemble d'un des domaines où l'extrapolation du concept linguistique d'énonciation a fait l'objet de discussions approfondies, où le transfert de la linguistique à la sémiotique audiovisuelle a soulevé maintes difficultés. En premier lieu le cinéma, qui a été conçu d'abord comme *langage*, « grammaire », code, etc., voire comme un système de signes, dont la définition des unités minimales n'a pas été facile<sup>91</sup>. De cette première étape, est symptomatique l'expression « parler cinéma », popularisée quelques années plus tard, et conçue comme la manière dont les « grands imagiers », les auteurs implicites et autres figures de la narratologie filmique, actualisent les possibilités immanentes d'une « langue », dans l'usage propre de la « parole filmique », à la manière saussurienne. À ceci près que la parole saussurienne, et surtout sous l'emprise de la « masse parlante », est bien plus qu'un usage : de multiples effectuations, des myriades d'exécutions continues dans le temps incompressible de l'action, soumise à des forces sociales, psychiques et sémiologiques. Quoi qu'il en soit, dans cette dualité langue / parole, la parole filmique revient alors à un énonciateur (le « sujet parlant » ou la « masse parlante »).

Peu après, le film a été vu comme un *texte*, une unité de sens plus vaste (par rapport au « signe »), un « ensemble signifiant » défini comme un univers semi-fermé, en dialogue avec l'institution cinématographique (entendue comme l'ensemble de tous les films, selon l'idée de Christian Metz), à laquelle il appartient. Dans cette étape, les « films concrets »<sup>92</sup> (les produits des énonciations) sont conçus comme le résultat d'une « écriture »<sup>93</sup>, un travail de production

---

<sup>91</sup> Même si la conclusion du fondateur de la sémiotique du cinéma, Christian Metz, à cette époque-là a été que le cinéma était un « langage sans langue », au début les discussions ont tourné autour des « articulations » d'un tel langage, sa « syntaxe » et, en général, sa « grammaire ». Voir, particulièrement (METZ, 1964).

<sup>92</sup> C'est-à-dire, simplement, « les films pris (un à un) », chacun avec une organisation particulière de thèmes et motifs, et un « usage singulier qui y est fait des divers procédés cinématographiques », des « relations qui unissent cet usage au « contenu » du film, etc. » (METZ, 1971 : 53).

<sup>93</sup> « On peut donc dire, en termes barthésiens, que le langage cinématographique a plus de points communs avec une écriture qu'avec une langue. Ce sont les cinéastes, non la population globale, qui ont fait le cinéma, comme ce sont les écrivains qui ont fait l'écriture. La langue appartient à tout le

textuelle qui doit être interprété selon ses « logiques internes », et surtout, selon la tension entre systèmes (singuliers) et codes (généraux). C'est également l'époque de l'introduction des modèles communicationnels (des émetteurs et récepteurs) et sa traduction sémiotique dans les actants de l'énonciation : énonciateurs et énonciataires. Mais on sait bien aujourd'hui qu'un « film », pour être une effectuation « concrète », est plus qu'un texte, à moins d'englober au moins dans le texte en question un écran et un espace de projection et d'écoute, et qu'en matière de « communication » le film propose et transmet plutôt des positions et des rôles à adopter et à adapter.

Comme pour anticiper cette objection d'aujourd'hui, et simultanément (et avec des développements postérieurs) le mariage sémiologie – psychanalyse a donné lieu à cette même époque à la théorie du *dispositif*, moitié description d'un « medium » technique, moitié description d'une expérience (pré)configurée dans l'espace de projection, et portée par une disposition psychique du spectateur. Dans cette étape la bascule commence à pencher du côté de l'énonciataire (« réel » ou « simulé », selon la perspective théorique), après l'intérêt exclusif de la première narratologie pour l'énonciateur : il s'agit de décrire, au moins au niveau général, l'expérience des êtres humains réunis dans une salle de projection, qui s'apparente aux états de rêve ou d'hypnose<sup>94</sup>. Le mouvement de bascule qui substitue ainsi l'énonciataire à l'énonciateur apporte un complément, voire une compensation, mais ne résout pas pour autant la question sous-jacente, à savoir celle de leurs relations, interactions et échanges de positions.

Finalement, l'intérêt pour les énonciataires a donné lieu à différentes théories de l'expérience de réception et d'interprétation filmiques. Par exemple, Michel Chion a proposé le terme *audiovision*, défini comme une *attitude perceptive*, une *activité* déployée par le spectateur. Les médias audiovisuels, nous dit Chion, « ne s'adressent pas seulement à l'œil »<sup>95</sup>. Et ajoute : « Ils suscitent chez leur spectateur –leur audio-spectateur– une *attitude perceptive* spécifique, que nous proposons, dans cet ouvrage, d'appeler l'*audiovision*. Une

---

monde, le cinéma et l'écriture sont choses de « spécialistes », bien qu'ils ne se confondent pas avec les idiolectes propres à chacun d'entre eux. C'est pourquoi le langage cinématographique évolue considérablement plus vite qu'une langue, et que son évolution se déploie dans une temporalité qui n'est pas de même échelle que celle des diachronies linguistiques, et rappelle plutôt celle des diachronies littéraires ». (METZ, 1971 : p. 201).

<sup>94</sup> Voir BAUDRY (1975) ; METZ (1975) et, plus récemment, BELLOUR (2009).

<sup>95</sup> CHION (2008 : p. 3).

*activité* qui, bizarrement, n'est jamais envisagée dans sa nouveauté ». Dans tout le livre que cet auteur consacre au concept d'audiovision, seules quelques lignes (dans l'introduction) sont dédiées à la définition du terme. Dans un autre ouvrage, Chion propose un glossaire, où on peut trouver une telle définition : le terme audiovision « désigne le type de perception propre au cinéma et à la télévision, mais souvent aussi vécue *in situ*, dans lequel l'image est le foyer conscient de l'attention, mais où le son apporte à tout moment une série d'effets, de sensations, de significations qui souvent, par un phénomène de projection dit *valeur ajoutée* sont portés au compte de l'image et semblent se dégager naturellement de celle-ci »<sup>96</sup>. Il s'agit donc d'un « type de perception », une « activité perceptive ». Et, bien que cette activité soit activée par les images à l'écran et par les sons dans les hautparleurs, il s'agit de l'activité du spectateur. Le mouvement de bascule s'accroît, puisqu'il est ici utile de se demander s'il s'agit toujours d'un énonciataire, instance reconstruite à partir du film, ou d'un spectateur « réel » (même s'il reste générique).

Parallèlement, chez les théoriciens « cognitivistes » du cinéma, la question ne se pose plus : ils ont également privilégié le spectateur, mais bien le spectateur réel, en essayant de modéliser son activité interprétative selon les paramètres de la psychologie et les théories de la perception. Plus que les « textes filmiques », ce qui intéresse ici ce sont les « inférences », les « hypothèses » et les réactions des spectateurs, résultats de leur activité corporelle et neuronale. Dès lors, s'agissant des spectateurs « réels » dotés d'un psychisme, d'organes sensoriels et d'activités neurocognitives, la question de l'énonciation est mise de côté, ou même rejetée comme non pertinente<sup>97</sup>.

Dans cette brève histoire, le concept d'énonciation apparaît de façon intermittente, et plus ou moins explicite. Ce concept a été lié à ceux de narration, de monstration, de "point de vue" et d'autres notions qui ont à voir avec le fait que le cinéma est un médium qui raconte, à travers des images et des sons, ainsi que des mots. Néanmoins, bien qu'il s'agisse d'un médium impliquant la perception, la dimension corporelle de l'expérience audiovisuelle a rarement - ou insuffisamment- été traitée. Nous voudrions donc aborder ici les problématiques

---

<sup>96</sup> CHION (2003 : p. 413).

<sup>97</sup> C'est principalement David Bordwell qui a questionné la pertinence de la théorie de l'énonciation appliquée au cinéma (BORDWELL, 1985). Cette position a été discutée, entre autres, par Jost et Gaudreault, en particulier en ce qui concerne la narration (GAUDREULT & JOST, 1990). Voir aussi (JOST, 1992, p. 28). Ces auteurs accusent Bordwell de tomber dans l'animisme, en affirmant que c'est « l'histoire » qui « se donne » au spectateur.

liées à la question de l'énonciation dans le cadre de la sémiotique du discours et de ses derniers développements : la sémiotique du corps et la théorie des pratiques sémiotiques qui proposent un cadre théorique de pertinence pour aborder la relation entre le corps et l'énonciation dans les dispositifs audiovisuels et les expériences qui s'y déroulent. Le rapide panorama qui précède met en évidence la nature du défi à relever : passer par la sémiotique des pratiques et la sémiotique du corps sans pour autant sortir du champ de l'énonciation !

### **Narration, monstration et énonciation**

À la différence de la littérature, il est indéniable que les « textes filmiques » activent des expériences perceptives réelles, « en direct », voire « primaires », à partir desquelles le spectateur doit construire un récit, une diégèse, un *monde possible*<sup>98</sup>. Et ce spectateur a un corps ou, mieux, *est un corps*. Cette question n'est pas passée complètement inaperçue dans les théories de l'énonciation audiovisuelle. Pour ne citer que quelques exemples, on peut mentionner Francesco Casetti, Gianfranco Bettetini, André Gaudreault, Roger Odin et François Jost.

Dans son ouvrage, *D'un regard l'autre : le film et son spectateur* Francesco Casetti s'efforce de décrire l'activité du spectateur sans sortir d'un cadre structuraliste et énonciatif. Il parle d'une solidarité entre l' « interlocuteur qui est anticipé par les images et les sons et l'interlocuteur dont le regard croise les regards venant de l'écran », bien qu'entre eux il y ait aussi des différences. Le corps de cet « interlocuteur » est, donc :

*...le support et la réserve d'un rôle. Le support avant tout. Si les lignes de force qu'un texte filmique s'emploie à mettre en place, et qui définissent à leur tour les articulations fondamentales du texte, visent d'un côté à transformer des hypothèses en une opération effective et de l'autre à donner une certaine stabilité à des actions en partie variables, elles*

---

<sup>98</sup> Il y a une forte tendance à privilégier le composant narratif par rapport au composant visuel du cinéma. Chez le premier Metz, par exemple : « La dialectique histoire/image le règne de l' « histoire » va si loin que l'image, instance que l'on dit constitutive du cinéma, s'efface à en croire certaines analyses derrière l'intrigue qu'elle a elle-même tissée, et que le cinéma n'est plus qu'en théorie art des images. Le film, que l'on croirait susceptible de donner lieu à une lecture transversale, par l'exploration à loisir du contenu visuel de chaque « plan », est presque à tout coup l'objet d'une lecture longitudinale, précipitée, déphasée vers l'avant et anxieuse de la « suite ». La séquence n'additionne pas les « plans », elle les supprime » (METZ, 1964 : pp. 62 – 63).

*doivent s'appuyer sur quelque chose ou sur quelqu'un. En d'autres termes, une présence effective est nécessaire pour ajuster des lignes de conduite dans une action et pour agencer des points de vue perpétuellement fuyants autour d'un point fixe. Dans ce sens, un corps, face à un rôle, est comme une plaque photographique qui révèle avec précision les contours d'une image et qui, en même temps, fixe le mouvement dans une pose canonique : donc comme une surface qui se grave et qui immobilise. En plus d'être un support, c'est, comme nous l'avons dit, une réserve. La plaque est aussi un palimpseste où chaque inscription se superpose aux traces précédentes et où chaque trace demeurée devient l'occasion d'un souvenir. En effet, les lignes de force du texte, bien loin de rencontrer des présences inertes, puisent en quelque chose ou en quelqu'un de quoi les alimenter et les informer ; elles rencontrent le réceptacle d'une compétence. (CASSETTI, 1990 : pp. 70 – 71)*

Curieuse métaphore pour désigner un corps vivant, celle de la *plaque photographique*, comme si l'objet inorganique était le corps et pas le film... Peut-être le même auteur a-t-il noté la difficulté de sa métaphore et n'a pas hésité à en proposer une autre, plus « interactive » : « dans ce sens-là, un corps, face à un rôle, est comme un registre permettant d'établir le passif et l'actif d'un compte ; comme l'avaliseur d'un prêt, capable d'interrompre les frais s'il est à découvert, ou de le relancer si on lui paie ses intérêts » (CASSETTI, 1990 : p. 71). En tout cas, la question de fond apparaît ici : si l'on cherche à situer le rôle du corps dans l'énonciation audiovisuelle, il faut également interroger l'« intercorporéité », l'échange et la transmission entre les deux corps énonçants, celui du pôle énonciateur et celui du pôle énonciataire. Casetti semble envisager cette interaction sur le mode socio-économique : un compte ouvert par l'un, un don provisoire (un prêt), appelant une sorte de contre-don de l'autre.

Dans les théories du cinéma venant d'Italie, l'autre auteur qui a traité plus en détail la question du corps dans l'énonciation audiovisuelle est Gianfranco Bettetini, qui a résolu l'opposition entre le « corps textuel » et le « corps réel » de l'énonciataire avec le concept de « prothèse symbolique ». D'après Bettetini il n'y a pas contiguïté entre le corps du spectateur et les signes dans l'écran :

*L'effet illusoire des signes filmiques satisfait l'activité de ses sens, l'organicité matérielle de son corps ; mais la connaissance de son immatérialité, de son manque de concrétion, fendille cette satisfaction et lance son activité psychique et, même plus, l'ordre du système symbolique inscrit aussi dans son corps (...). Le corps du spectateur est reproduit symboliquement dans un simulacre tendanciellement homogène, similaire à celle incorporée dans le texte par le*



*sujet énonciateur et révèle, dans cet acte, toute sa nature relationnelle (...) Le corps du spectateur "prolonge" son action en devenant un "autre" corps de nature symbolique, en se construisant une véritable et propre prothèse symbolique. (BETTETINI, 1996, pp. 34 - 35)<sup>99</sup>*

Bettetini comprend la prothèse dans un sens large, comme un dispositif qui étend la capacité d'action d'un organe en permettant de pénétrer des lieux exclus des possibilités naturelles du corps. Dans le cas du cinéma, il n'y a pas besoin de prothèses visuelles ou auriculaires proprement dites –puis les images et les sons arrivent directement aux organes sensoriels correspondants chez le spectateur dit « normal »-. Le sens du terme « prothèse » est ici particulier : l'image audiovisuelle s'impose au destinataire comme un monde apparemment ouvert à l'interaction corporelle ; néanmoins, le spectateur se rend rapidement compte que cette interaction est incomplète et illusoire : il manque la poli-sensorialité de l'expérience réelle (olfaction, gustation, etc.) et la spatialisation. Le spectateur, donc, « finit par se sentir auteur, en remplaçant l'impuissance sensorielle avec un dispositif symbolique qui remplace l'action des sens non stimulés : justement, avec une prothèse symbolique » (35)<sup>100</sup>.

Cela nous renvoie au champ de l'énonciation énoncée : autour du « vu » et de l'« entendu » par l'énonciateur le simulacre de l'énonciataire est construit, et la prothèse symbolique du corps du destinataire (réel) interagit avec ce simulacre :

*...on peut dire que tout le corps du destinataire se prolonge dans une prothèse symbolique qui est poussée à souligner, en vertu du projet énonciatif immanent au texte, le simulacre symbolique du sujet de l'énonciation : les yeux, les oreilles, la synesthésie du corps absent devient les organes sensoriels du corps symbolique produit par le spectateur, avec tous les effets d'illusion et du feed-back sensoriel demandés au corps psychique par le manque des parties du corps physique et sa substitution avec la prothèse (p. 36).*

De cela Bettetini conclut que le corps du sujet de l'énonciation peut être interprété sous la perspective d'un « savoir organisé », un ensemble d'informations pré-construites qui s'articulent dans le texte comme acte discursif. L'auteur parle d'une « encyclopédie

---

<sup>99</sup> Pour la version originelle, voir (BETTETINI, 1979).

<sup>100</sup> Dans une œuvre plus récente, Bettetini s'occupe aussi du composant sonore des médias audiovisuels : *Il timpano dell'occhio. Gli intrecci e i giochi dei suoni e delle immagini* (BETTETINI, 2009).



organique » qui s'offre à l'énonciataire, qui organise aussi sa propre structure symbolique, laquelle est née de la rencontre du savoir de l'énonciateur avec le sien. Le spectateur « voit le film et en même temps il l'aide à être texte, à être acte communicatif » (37).

Ainsi, le corps réel disparaît à nouveau sous l'ombre du simulacre : « l'échange est fermé dans un lieu immatériel où les corps sont absents et où une connaissance est produite, une acquisition du savoir, une incorporation symbolique... » (37). En tout cas, tant Casetti que Bettetini sont sensibles au tournant dit « pragmatique » des années quatre-vingt, qui a permis de commencer à parler de l'énonciation « en acte ». Un tournant qui met l'accent sur l'activité du spectateur et son expérience perceptive.

Dans les théories de l'énonciation depuis les années quatre-vingt, la question du corps a été cantonnée, au rôle discriminant des deux types de sensations sollicitées, en particulier depuis qu'on a commencé à faire la distinction entre la *focalisation* (relative au *savoir*) et l'*ocularisation* et l'*auricularisation*<sup>101</sup>, rapportées à *voir* et *entendre* ; c'est-à-dire, la distinction entre un faire cognitif et des faires *perceptifs* (JOST, 1987) ; (JOST, 1992) ; (GAUDREULT & JOST, 1990).

On peut retrouver la distinction entre montrer et raconter (*showing and telling*, dans la tradition anglaise), comme l'a fait André Gaudreault (Gaudreault, 1999), jusque dans la distinction entre mimésis et diègèsis de la tradition grecque. Et bien que pour ce théoricien la distinction la plus précise doit être entre *diègèsis mimétique* (comprise comme représentation scénique) et *diègèsis non mimétique* (la narration « directe » du « poète »), en réservant le terme *mimésis* pour désigner la représentation poétique en général, la dichotomie *narration vs monstration*, valable pour le théâtre et le cinéma, est une des catégories les plus importantes dans son approche narratologique<sup>102</sup>. En ce qui concerne l'énonciation visuelle (pour le théâtre, puis pour le cinéma), le terme *monstration* est le plus intéressant, car il dépasse la

---

<sup>101</sup> Dans son ouvrage *Les Espaces subjectifs* (sur lequel nous reviendrons), Jacques Fontanille met en question la pertinence de ces distinctions. Nous reviendrons plus tard sur ce sujet.

<sup>102</sup> D'après Gaudreault, il faut distinguer la *narration scripturale* de la *monstration scénique*. C'est dans la première qu'on peut utiliser le *showing* et le *telling* de la narration scripturale, tandis que dans la deuxième on utilise la distinction *diègèsis mimétique / diègèsis non mimétique* (GAUDREULT, 1999 : pp. 86 – 89).

discrimination sensitive. Chez Gaudreault, ce terme remplace le terme moins précis de « représentation » :

*...j'ai suggéré, il y a quelque temps, l'utilisation systématique d'un terme, « monstration », déjà existant<sup>103</sup> (quoiqu'il n'ait pas encore fait son entrée dans les dictionnaires) et qui commence à s'imposer, pour caractériser et identifier ce mode de communication d'une histoire qui consiste à montrer des personnages qui agissent, plutôt qu'à dire les péripéties qu'ils subissent, et pour remplacer ce terme trop marqué, trop galvaudé et par trop polysémique qu'est « représentation ». Le choix du terme de « monstration » m'apparaît aujourd'hui d'autant plus heureux qu'il m'a permis à l'époque de former le substantif « monstrateur », que je propose à nouveau ici pour identifier et singulariser cette entité théorique, équivalente du narrateur fondamental du scriptural, qui serait responsable de la communication du récit scénique (GAUDREULT, 1999 : p. 87).*

Dans le cas du cinéma, Gaudreault associe le travail du monstateur à l'activité du tournage, et donc à l'appareil de prise de vues. Ce travail est paradoxal si on pense que le temps du tournage est le présent, défini par Benveniste comme « la coïncidence de l'événement décrit avec l'instance du discours qui le décrit », tandis que dans le temps de la projection (c'est-à-dire, le temps du spectateur) « l'événement décrit » par l'appareil est déjà passé : « L'illusion de présent que procure la vision du plan n'est donc, en définitive, qu'un simulacre de présent » (GAUDREULT, 1999 : p. 105). En d'autres termes, d'après Gaudreault, le monstateur filmique opère toujours un débrayage temporel. Mais loin de résoudre le paradoxe temporel que provoque la désynchronisation entre les deux pôles de l'énonciation, Gaudreault conclut en complexifiant les instances de la narration :

*Le récit filmique serait, dans son ensemble, le produit de la superposition de deux couches distinctes de narrativité, chacune d'elles étant due respectivement à chacune des deux « articulations » (au sens non linguistique) de la double mobilité caractéristique du cinéma : l'articulation photogramme et photogramme, et l'articulation entre plan et plan. (...) Pour produire un récit filmique pluriponctuel, il faut d'abord faire appel à un monstateur qui, lors du tournage, « met en boîte » une multitude de micro-récits (les plans) qui ont tous et chacun, en définitive, une certaine autonomie narrative. (...) Il faut ensuite faire appel à une autre instance narrative, le narrateur, qui, se saisissant de ces micro-récits, pourra, s'il est*

---

<sup>103</sup> Selon Gaudreault, le terme a été utilisé avant par Betty Rojzman en 1981, dans le champ du théâtre, et par Xavier De France pour le cinéma, en 1982 (Gaudreault l'a utilisée par première fois dans sa thèse de doctorat en 1983).

« *narrativement* » euphorique, en travailler la substance narrative pour annuler l'autonomie des plans produits par le monstateur et y inscrire le parcours d'une lecture continue, consécutive au regard qu'il aura posé sur cette substance et qu'il y aura transposé. (GAUDREULT, 1999 : p. 107)

Pour différencier le monstateur filmique du monstateur du récit scénique, Gaudreault fait une autre distinction entre le « monstateur profilmique », responsable de la mise en scène, et le « monstateur filmographique », responsable de la mise en cadre. Au-dessus d'eux il place un « méga-monstateur filmique », responsable de la « mise sur film », et qui travaille parallèlement avec le « narrateur filmographique », en charge de la « mise en chaîne ». Et par-dessus toutes ces instances, il y a un « méga-narrateur filmique », le « grand imagier » de Laffay, responsable de la « mise en film ». L'ensemble de ces instances configure le *dispositif énonciatif du récit filmique* selon André Gaudreault. La prolifération des instances narratives commence, mais ne résout toujours pas l'énigme du rapport entre énonciateur et énonciataire.

La proposition de Gaudreault a été reprise et développée par Roger Odin, qui a ajouté encore d'autres instances pour résoudre quelques problèmes induits par ce modèle. Odin propose ainsi un *narrativiseur*, responsable de la narrativisation, entendu comme le procès que « mobilise une modalité de production de sens assez particulière : j'en reste à une sorte de *pré-compréhension*. Au niveau de la première sous-couche narrative (celle des micro-actions repérables à l'intérieur d'un plan), le fil est appréhendé comme une poussière de sens, sur un mode un peu analogue à ce qui se passe dans l'expérience quotidienne, lorsque je regarde bouger les choses autour de moi ; au cinéma, la relation entre mouvement et illusion de réalité se fait d'ailleurs à ce niveau » (ODIN, 2000 : p. 29). Le « Je » en question est-il celui de l'énonciataire ? Sans aucun doute, puisqu'il semble difficile de prêter à l'énonciateur cette activité de « précompréhension » calquée sur l'expérience quotidienne. Roger Odin ne se contente donc pas de complexifier la hiérarchie des instances d'énonciateur: il commence à s'intéresser à leurs interactions avec l'énonciataire.

Du côté des instances « responsables » de l'énonciation filmique, il y aura deux niveaux de narrativité selon Odin : le premier est sous la responsabilité du *narrativiseur* qui produit « un sentiment de narrativité ». Le second niveau, celui de la production du récit proprement dit, est divisé en deux : le récit en un seul plan, dépendant du monstateur, et le récit en plusieurs plans qui est sous la responsabilité du narrateur (et qui régit le travail du monstateur). Finalement, Odin ajoute aussi un dernier niveau, celui de la *discursivisation*,

entendue comme « la structuration paradigmatique de ces actions en fonction des valeurs qu'elles véhiculent » (ODIN, 2000 : p. 35). L'auteur comprend ce niveau comme l'étape des « structures profondes » du parcours génératif de la sémiotique d'A.J. Greimas. Les autres étapes décrites par lui sont équivalentes aux étapes postérieures de ce parcours : « tout récit finit de la sorte en *discours*, et toute intervention d'un *narrateur* se double de la construction d'un *responsable du discours*. Il arrive qu'en conclusion du récit, ce responsable se manifeste explicitement et résume la signification du discours en tirant la morale de l'histoire »<sup>104</sup> (ODIN, 2000 : p. 36).

Mais Roger Odin ajoute aussi un autre concept de grand intérêt pour une théorie de l'énonciation qui prend en compte la dimension corporelle et somatique, notamment dans l'interaction avec l'énonciataire : la *mise en phase*, entendue comme « le processus qui me conduit à *vibrer au rythme* de ce que le film me donne à voir et à entendre. La mise en phase est une modalité de la participation affective du spectateur au film » (ODIN, 2000 : p. 38). A regarder attentivement cette participation du spectateur, on peut se demander s'il s'agit d'un énonciataire formel ou du « vrai » spectateur de « chair en os » ? Il est d'abord question du rapport entre la *mise en phase*, la narration et l'énonciation. En effet, Odin parle de la *mise en phase narrative*, c'est-à-dire, « le processus qui me conduit à *vibrer au rythme des événements racontés*. Familièrement, on désigne ce processus en disant que l'on a « marché » » (ODIN, 2000 : p. 39). Admettons pour l'instant que ce spectateur soit toujours intégré au champ de l'énonciation, par « immersion » corporelle.

La contrepartie de ce processus d'immersion c'est l'effet de *déphasage* qui peut prendre plusieurs formes : déphasage « plastique », actoriel, monstratif, interne ou généralisé.

---

<sup>104</sup> Dans sa contribution au numéro 38 de *Communications* (1983), dédié à « l'énonciation au cinéma », Jacques Aumont propose quatre sens de la locution « point de vue » : 1. L'emplacement depuis lequel on regarde et, pour le cinéma, l'emplacement de la caméra relative à l'objet regardé. 2. La vue elle-même, en tant que prise depuis un certain point de vue. Dans le cinéma, à ce niveau correspond le problème de l'« encadrement ». 3. La correspondance de ce dernier avec un « point de vue » narratif, quand, par exemple, le cadre est représentation d'un regard, « celui de l'auteur ou celui du personnage ». 4. « Le tout, enfin, est surdéterminé par une attitude mentale (intellectuelle, morale, politique, etc.) qui traduit le jugement du narrateur sur l'événement ». Aumont qualifie ce point de vue là comme « prédicatif » (Aumont, *Le point de vue*, 1983 : pp. 4 – 5). On peut assimiler cette proposition avec les différents niveaux de Gaudreault et aussi d'Odin, le « point de vue prédicatif » étant le résultat de la discoursivisation.

Il s'agit des moments où le spectateur se détache de l'histoire et prend conscience des aspects énonciatifs qui, dans la mise en phase narrative, restent invisibles. Mais « toute autonomisation d'un élément ne provoque pas automatiquement un effet de déphasage ; si la lecture du film fait jouer cette autonomisation au profit du récit, certains écarts peuvent être récupérés au bénéfice de la mise en phase » (ODIN, 2000 : p. 43). La mise en phase énonciative, donc, « mobilise *sans les effacer* certaines marques d'énonciation en tant qu'*adjuvants narratifs* » (p. 44).

Selon cette proposition, la monstration n'est pas le seul élément qui mobilise la perception et le corps (ou l'appareil de vision du corps) : la mise en phase implique aussi la composante somatique du spectateur, qui *vibre* avec le récit ou même avec les composants de la structure énonciative. Pourtant, dans ce modèle, la façon dont les éléments cognitifs liés à l'information narrative, et les aspects pragmatiques de l'action dans le récit interagissent avec les aspects thymiques ou somatiques d'une telle « vibration » spectatorielle, reste inexplicite. En outre, la monstration dont parlent Gaudreault, Odin et d'autres théoriciens semble inclure tous les aspects de la mise en scène audiovisuelle, mais pas clairement le rôle des éléments non visuels (la bande-son), et de la dimension corporelle dans toute sa complexité (ce qu'on va appeler plus loin les *modes du sensible*). Il y a donc enfin un corps, mais son rôle est encore plus ou moins confus, et surtout instable.

Dans son ouvrage écrit avec François Jost, Gaudreault reprend la question de la monstration et, cette fois, les auteurs consacrent quelques lignes au « double récit » du cinéma sonore : le récit visuel en charge de la monstration, et le récit sonore, construit par les paroles et les sons en général. Normalement, nous disent Gaudreault et Jost, tout est fait pour que le dialogue et les voix réduisent les ambiguïtés des énoncés visuels et pour qu'on ne perçoive pas cette dualité du film sonore. Il y a aussi des cas dont le récit A (visuel) et le récit B (sonore) s'opposent ou sont en contraste. Mais ce qui importe ici c'est que, en tant qu'il existe un double récit, il doit y avoir aussi différentes entités responsables de ces énonciations. En suivant André Gardies, les auteurs proposent trois sous-énonciateurs, chacun responsable, respectivement, de l'iconique, du verbal, et du musical. Néanmoins, ces énonciateurs travaillent sous la baguette d'un « méga-narrateur filmique », le grand imagier d'Albert Laffay.

Un développement plus récent de cette ligne théorique met en évidence la complexité d'une telle proposition. Dans son ouvrage *La fiction au cinéma*, Allain Boillat parle des

*instances diégétiques déléguées* par rapport au « méga-énonciateur » filmique : il y aura donc des monstateurs, mais aussi des « scripteurs », des « partiteurs » ou des « bruiteurs »... (Boillat, 2006). Et la multiplication des « instances énonciatives » peut continuer jusqu'à construire des simulacres formels de toute l'équipe de production cinématographique. Une telle prolifération d'instances énonciatives jette le doute sur l'économie théorique et explicative d'un tel modèle. Et surtout, la fuite en avant laisse dans l'ombre (et même en l'épaississant) l'interaction entre les deux pôles de l'énonciation.

### **Focalisation, ocularisation, auricularisation. Le « point de vue » revisité.**

François Jost a développé de façon systématique les concepts d'*ocularisation et auricularisation*. L'auteur commence par souligner la nécessité d'opposer ces concepts à celui de *focalisation* : « Pour définir le concept de narration, il faut donc tracer la frontière entre raconter et voir. Méthodologiquement, il n'est pas possible d'assimiler purement et simplement la question qui voit? à ce qu'on nomme, en théorie littéraire, la focalisation. Si pour les théoriciens de la littérature ce concept recouvre principalement les problèmes de savoir narratif, pour ceux du cinéma, à l'inverse, il désigne plutôt « la question proprement narratologique de savoir quel est le foyer de perception (...) choisi pour médiatiser la représentation » (JOST, 1987 : pp. 21 – 22).

Jost a proposé le terme d'*ocularisation* qui a connu une belle carrière dans le monde de la narratologie filmique. Si la focalisation désigne ce que sait un personnage, l'ocularisation parle de « ce que la caméra montre et ce que le héros est censé voir ». Néanmoins, l'ocularisation se comporte de façon similaire à la focalisation, dans la mesure où il faut en distinguer différents types en fonction de l'entité narrative qui la porte. Le critère qui permet de différencier cette position est précisément le *corps* : l'ocularisation sera interne, lorsque la caméra prend « la place de l'œil du personnage » ; mais cette ocularisation interne peut être primaire « dans le cas où se marque dans le signifiant la matérialité d'un corps ou la présence d'un œil qui permet immédiatement, sans le secours du contexte, d'identifier un personnage absent de l'image » (pp. 27 – 28). Si la subjectivité d'une image « est construite par le montage, les raccords ou par le verbal », c'est-à-dire, par une contextualisation (et pas par un positionnement corporel), l'ocularisation interne sera *secondaire*<sup>105</sup>.

---

<sup>105</sup> François Jost décrit aussi des cas où la prise de vue est ambiguë par rapport à un personnage ou à la narration verbale : « quelle place accorder dans notre typologie à tous ces moments où nous voyons en un éclair, ce que le personnage (narrateur) imagine, se rappelle, crois voir, etc., en bref ce que l'on

L'ocularisation zéro correspond au cas où l'appareil est placé « *en dehors* » du personnage. Entendue comme le résultat des prises de vue « neutres », l'ocularisation zéro doit être différenciée des cas où « la mise en scène ou en cadre a plutôt pour fonction de livrer au spectateur une information que ne possède pas forcément le héros : l'ocularisation prend valeur de focalisation ». L'auteur considère que, dans ces cas où « le spectateur ne partage son point de vue avec aucun personnage et en tire un avantage cognitif, on pourrait presque parler d'ocularisation spectatorielle » (JOST, 1987 : p. 28). Il y a aussi un autre cas qui implique le spectateur : le « regard à la caméra », qui suppose le franchissement de la frontière du monde diégétique. Ainsi, ce n'est pas seulement le corps du personnage qui est convoqué, mais aussi le corps spectatorial. Mais il est étrange que la place du spectateur ne soit évoquée que parce l'ocularisation est neutralisée : on pourrait tout aussi bien considérer les ocularisations internes comme des « propositions » faites au spectateur, si ce n'est des manipulations de sa position qui lui seraient imposées, voire des provocations qui lui seraient opposées.

Il n'en reste pas moins que le concept d'ocularisation permet de convoquer à la fois l'énonciateur et l'énonciataire à travers un positionnement perceptif, grâce à la médiation corporelle et sensitive. Cette médiation corporelle est étendue à l'écoute, et le problème de la multiplication de critères et d'instances d'analyse apparaît une fois plus. Néanmoins, le fait d'inclure la dimension sonore de la perception donne lieu à des réflexions intéressantes...

### **Auricularisation, « point d'écoute » : le corps audiovisuel**

*...Si on peut caractériser la relation du personnage à ce que montre la caméra, il semble tout autant nécessaire d'envisager la bande-son par rapport au même personnage. Il est bien évident que l'occurrence d'un bruit ou d'un dialogue, comme le mixage lui-même, peut s'ancrer de façon plus ou moins subjective dans la fiction. Par exemple, lorsque sur l'image d'un personnage qui ne parle pas, on entend une bribes de texte ou un bruit que le contexte spatial ne justifie pas, bien souvent il faut interpréter cette intervention sonore comme un*

---

nomme « images mentales » ? » (Jost, 1987 : p. 31). L'auteur introduit ici des *opérateurs de modalisation* (des éléments techniques du montage, filtrage, etc.) qui permettront de parler des ocularisations modalisées. Quand ces opérateurs disparaissent, « tout se passe comme si le personnage confondait son souvenir, son rêve, son imagination, avec une perception ». Jost propose dénommer cette ocularisation comme « interne hallucinée » (p.32). Une fois plus c'est responsabilité du spectateur de distinguer entre une ocularisation et une image mentale, à travers des « inférences sémantiques ».



*souvenir ou une « imagination » de ce personnage. Il existe des cas plus complexes où le mixage crée de véritables points de vue sonores [auxquels] je donne le nom d'auricularisation.* (JOST, 1987 : p. 23)

Voici la première définition de l'auricularisation selon François Jost. Pourtant, l'auteur nous met en garde contre les dangers d'une simple transposition des catégories de l'ocularisation au monde de l'écoute : il ne suffirait pas « d'un simple ajustement de la théorie du point de vue visuel pour dégager les concepts opératoires dans une analyse de la bande sonore. Ce serait compter sans la résistance que le matériau oppose à toute entreprise narratologique pure et désincarnée »<sup>106</sup>. (p. 45). Le premier obstacle pour en faire une telle transposition c'est « le caractère acousmatique de l'écoute », ce qui est le cas lorsqu'il s'agit de sons enregistrés. Et dans le cas du cinéma, comme bien l'a dit Roger Odin, l'écoute est doublement acousmatique : « par rapport aux sources directes des sons (il s'agit de sons enregistrés) ; et par rapport aux sources indirectes : le spectateur ne voit pas le haut-parleur. Ainsi, radicalement déconnecté de son origine, le son filmique flotte dans l'espace de projection » (JOST, 1987 : p.46).

Nous sommes donc bien dans la perspective du spectateur. De cette condition « acousmatique » du son au cinéma, Jost tire les conséquences : le son n'est pas forcément *relayé* par l'image (il peut avoir sa source en dehors du monde de la diégèse) ; le son n'est pas latéralisé (en exceptant la stéréophonie) ; et le son n'est pas localisé *a priori* dans une source visuelle, mais sa localisation aboutit, pour le spectateur, à un *ancrage* de sa source apparente dans un objet visualisé ou non. Les conditions de la spatialité des sons dans les salles de cinéma (et dans les *Home Theater* et autres systèmes sonores domestiques) ont changé radicalement, au point que la distribution tridimensionnelle des sons n'est plus une exception négligeable. La description de Jost doit être aujourd'hui révisée : la latéralisation des sons est de plus en plus contrôlée par les réalisateurs avec des intentions narratives et/ou esthétiques, notamment leur ancrage dans les éléments de l'univers diégétique, et pour capter et manipuler l'attention auditive du spectateur. C'est pourquoi la question de la spatialité des sons et de leurs ancrages en relation avec les images et leur contexte (principalement le hors-champ) nous fait penser qu'il semble nécessaire de reconsidérer les conditions perceptives de réception des sons, dans ce qu'on va appeler plus tard le « dispositif audiovisuel ».

---

<sup>106</sup> Pourtant, l'auteur parle aussi d'auricularisation *interne (primaire et secondaire), zéro, interne modalisée*, etc., avec les mêmes adjectifs qui qualifient l'ocularisation.



Pour l'instant, ce qui intéresse la narratologie ce sont les relations entre les sons émis pour le spectateur, et les éléments (objets, ambiances, personnages) de l'univers diégétique. Étrangement, le récit visuel serait plus tourné vers l'énonciateur, et la mise en scène sonore, vers l'énonciataire, comme si l'histoire du médium cinématographique pesait encore sur la théorie du film. On peut se demander, donc, pour reprendre la question de l'auricularisation, en suivant Jost, « d'une part, si le personnage peut entendre ce que nous entendons, d'autre part, si ce que nous entendons est restreint à son champ auditif » (p. 47). Et si, pour le personnage, il est possible de latéraliser et de localiser un son.

À cet égard, il faut noter que la « vraisemblance » spatiale des sons ne dépend pas strictement de leur position réelle dans la projection et leur seule relation avec les « points de vue » des personnages, mais aussi de l'intelligibilité des informations perceptives. On trouve ici un problème particulier de « l'énonciation sonore », un conflit dont la solution n'est pas toujours claire. L'exemple où ce conflit apparaît clairement est celui des dialogues et de leur relation avec leurs « émetteurs » diégétiques. François Jost fait une longue citation d'un texte à l'usage des ingénieurs du son à cet égard :

*...au cours d'un dialogue entre trois personnages, vu par exemple de profil, puis de face, ensuite de loin, il ne saurait être question de déséquilibrer trop franchement ce dialogue à trois par un changement de tonalité ou par une variation de plan sonore. Si, par la suite, un des personnages prend congé et que la caméra souligne ce travelling de départ, il importe au contraire que le son respecte ce mouvement de plan et traduise un changement d'ambiance acoustique. Là, également, se pose la définition du point de fuite ; il est presque toujours situé à l'emplacement d'où le spectateur voit se dérouler la scène. (JOST, 1987 : p. 54)*

Dans des cas comme ceux-ci, il semble inévitable de privilégier le point de vue du spectateur : « soit on supprime l'idée même de point de vue, lorsqu'une servilité trop grande aux différentes ocularisations internes secondaires aurait pour résultat de 'déséquilibrer le dialogue', soit on se place du point de vue du spectateur ». À cet égard, on peut dire qu'il y a un conflit entre deux notions différentes de « point de vue sonore » ou, pour utiliser une expression moins contradictoire, de « point d'écoute ». Tout se passe comme si, d'une autre manière, la désynchronisation et la distension entre les deux pôles de l'énonciation étaient toujours à l'œuvre.

Selon Michel Chion, cette dernière expression désigne : « 1) le point à partir duquel il nous semble pouvoir dire que nous entendons un son comme proche ou lointain de nous, point qui soit concorde avec la place de la caméra, soit en est différent (cas fréquent du personnage éloigné dans l'image et de sa voix entendue proche). C'est alors le point d'écoute au sens *spatial*. 2) Le personnage par les oreilles duquel il est suggéré que nous entendons un son (par exemple, s'il parle au téléphone, et que nous entendons distinctement son interlocuteur comme lui-même est censé l'entendre). C'est alors le point d'écoute au sens *subjectif* » (CHION, 2008 : pp. 223 – 224).

Cette différence ne se confond pas avec celle qui oppose l'auricularisation interne et l'auricularisation zéro. L'idée d'un point d'écoute spatial implique que la description des sons au cinéma prenne en compte la scène de projection concrète et le placement « réel » du spectateur. Par exemple, Laurent Jullier préfère parler de « mondes » qui peuvent être habités par les éléments de la bande-son : « le *monde narratif*, les *hétéro-univers*, les *fosses*, et les *mondes de la fabrication et de la projection* (ou de la *diffusion* s'il s'agit d'une émission TV) ». Et Jullier continue : « on ne dira donc pas qu'un son est diégétique ou extra-diégétique, mais qu'il se situe quelque part le long de cet axe qui va du strict *monde narratif* (des quatre cités ci-dessus, c'est le plus proche de la traditionnelle catégorie du « diégétique ») au *monde de la projection*, le plus éloigné du « diégétique » (JULLIER, 1995 : pp. 103 – 104).

Or, bien qu'il soit vrai que dans le cinéma la frontière entre le « texte filmique » et la scène pratique où se déroule l'expérience d'« entendre et voir » un film n'est pas très claire, il y a ici un saut qualitatif qui mérite quelques précautions. La proposition de Jullier révèle un des grands problèmes des théories de l'énonciation audiovisuelle : d'une part, les théoriciens font de grands efforts pour différencier des instances et des figures des « énonciateurs » audiovisuels et d'autre part, la plupart des narratologues accordent pourtant, pour finir, la prééminence au spectateur, c'est-à-dire, à l'énonciataire « réel ». Dans cette perspective, les différentes instances « énonciatrices », conçues comme des simulacres textuels, sont converties en résultats d'inférences, d'hypothèses et autres « activités cognitives » des spectateurs.

Pourtant, il est possible d'opposer à cette tendance et à cette dérive le projet de modéliser les positionnements perceptifs de ces spectateurs, de manière à ce que la scène pratique de l'audio-visualisation devienne elle aussi une construction sémiotique à part

entière. La proposition de Jullier reste à cet égard très suggestive, car, pour rendre compte de l'écoute filmique et de la position proposée au spectateur, il est conduit à englober le *monde narratif* dans le *monde de la projection*, et à envisager de ce fait même un autre périmètre pertinent pour l'analyse, celui de la situation filmique, voire de la pratique cinématographique.

La sémiotique du discours et des pratiques offre, dans ses développements récents, des outils théoriques que permettent d'engager une telle construction. Il s'agit, comme nous l'avons annoncé dans l'introduction, de la sémiotique du corps et des modes du sensible, et de la théorie des pratiques sémiotiques et des niveaux de pertinence de l'analyse sémiotique. Pour encadrer la scène pratique de l'expérience audiovisuelle, nous reprendrons le concept de « dispositif », emprunté à la tradition des théories filmiques.

### **L'énonciation audiovisuelle et les modes du sensible**

On a vu que la modélisation de l'expérience perceptive audiovisuelle, exprimée avec des concepts comme « ocularisation » ou « auricularisation », permet d'aborder la question du corps dans la production de sens au cinéma. Il s'agit certes d'une réduction du corps à *l'œil et à l'oreille*, mais le fait de s'interroger sur les *positions du corps qui regarde et qui écoute*, et sur la manière dont ces positions gèrent l'accès à l'information narrative et thématique du film, donne déjà une longueur d'avance sur les théories narratologiques désincarnées développées à partir du modèle littéraire<sup>107</sup>. Le phénomène que nous sommes en train d'analyser, l'audiovision semble nous forcer à faire une telle réduction : n'est-ce pas à l'œil et à l'oreille que, de toute évidence, s'adressent les médias audiovisuels ? Mais si on généralise le rôle sémiotique du corps, même l'audiovision sera envisagée d'une autre façon. A ce stade de la réflexion, il faut reprendre la question de l'énonciation et les problèmes dérivés des relations entre le *savoir* et le *percevoir*.

Depuis Benveniste, l'énonciation relève de la question de la *subjectivité dans le langage*. Et ce n'est pas différent pour le cinéma. Néanmoins, dans ce cas les théoriciens ont pris différents points de référence pour développer une théorie de l'énonciation

---

<sup>107</sup> Jost a essayé aussi d'étendre la question sensorielle dans le cinéma aux autres sens, en particulier l'odeur et le toucher (JOST, 2014 : pp. 61 – 85).

cinématographique. Outre Benveniste<sup>108</sup>, on peut mentionner l'influence d'Oswald Ducrot et la notion d'« énonciation polyphonique », ainsi que la proposition narratologique de Gérard Genette et de ses épigones, qui ont développé les concepts de « mode », « point de vue » et « focalisation », mais aussi de « voix » et « narration », parmi d'autres, rapportés en différents degrés à la question de l'énonciation. On peut mentionner aussi la grande influence de l'ouvrage d'Umberto Eco, *Lector in Fabula*, et son introduction de « l'esthétique de la réception » dans le domaine de la sémiotique, ainsi que d'autres références théoriques (notamment, la théorie des « mondes possibles »), qui ont généré de nombreuses résonances dans la théorie du cinéma. Finalement, la théorie de l'énonciation issue de la tradition greimassienne a eu aussi des répercussions, bien qu'elle soit restée plus périphérique.

Après ces auteurs « fondamentaux », qui n'ont pas parlé directement ou exclusivement du cinéma, viennent les « sémiologues » et « narratologues » du cinéma qui ont essayé d'accommoder les théories linguistiques et discursives au domaine particulier des films et, ensuite, des médias audiovisuels en général. Jacques Fontanille se place dans cette perspective, notamment dans *Les espaces subjectifs. Introduction à la sémiotique de l'observateur* (FONTANILLE, 1989). Dans ce livre l'auteur essaie de proposer une théorie globale de l'observateur, avec un degré de généralité qui permet de la projeter sur tous les domaines discursifs, mais particulièrement, sur les textes verbaux (littéraires), et dans le domaine de la peinture et du cinéma. Son approche est ancrée dans la théorie de l'énonciation, et l'auteur précise :

*Celle-ci peut être entendue aussi comme la réalisation particulière des structures séminarratives –elle s'appellera alors « mise en discours » –que comme une interaction entre les sujets d'énonciation –elle s'appellera alors « communication » – (...) La sémiotique discursive s'occupe en fait de l'intersection de ces deux domaines : elle rend compte de ce qui, dans les formes mêmes de la mise en discours, et dans les opérations qui leur sont sous-jacentes, concerne, entre autres, l'interaction entre les sujets du discours (p. 6).*

Il s'agit donc de la « mise en discours » du savoir, liée aux questions de « perspective » et de « point de vue », manifestations d'une question plus générale : comment le sujet construit-il la signification ? Selon l'auteur, et dans cet ouvrage, il n'y a que deux réponses à cette question : (i) une approche cognitive, que fera appel aux opérations psycho

---

<sup>108</sup> Auteur qui est, dans une mesure plus ou moins grande, mentionné par presque tous les théoriciens du cinéma.

cognitives d'un sujet humain réel, et ii) une approche sémiotique, qui tentera d'établir, en interrogeant l'immanence du discours, le simulacre des opérations de l'énonciataire idéal.

Dans cette dichotomie on peut placer toutes les discussions sur l'énonciation au cinéma. D'un côté, les auteurs qui penchent pour le spectateur et en reconstituent les activités, que ce soit au sein de théories cognitives plus ou moins développées ou « radicales »<sup>109</sup>, ou bien d'un point de vue *pragmatique*, où la situation de communication est privilégiée : le spectateur est confronté au texte filmique en cherchant à en construire le sens au travers d'hypothèses, d'inférences et de modélisations portant sur les activités cognitives ou pragmatiques de ce(s) autre(s) être(s) humain(s) responsable(s) des images et de sons. Et d'un autre côté, les auteurs qui adoptent une position « immanentiste », et cherchent en revanche à construire et à décrire les « simulacres » textuels des énonciateurs et énonciataires. Notre grande question est comment placer l'audiovision dans cette double perspective ?

Dans *Les espaces subjectifs* Fontanille prend le deuxième chemin, dans la mesure où « seule l'analyse de l'énoncé, le repérage des significations immanentes possibles, l'identification de la signification sélectionnée et la description du mode de sélection sont adéquats dans cette seconde perspective » (p. 6).

Ainsi, Fontanille développe toute une sémiotique de l'observateur, avec plusieurs nouveautés théoriques : la description de la subjectivité non seulement en sa dimension cognitive, mais aussi en ses dimensions passionnelle (thymique) et pratique (ou pragmatique) ; la description des embrayages/débrayages de chacune de ces dimensions ; la configuration d'un réseau de concepts pour placer les instances subjectives dans le discours : focalisateur, spectateur, assistant, assistant-participant ; la construction des macro-actants de l'énonciation : l'informateur et l'observateur ; et tout un modèle de la modalisation cognitive de l'espace qui permettra de décrire la subjectivité au cinéma.

Néanmoins, cet ouvrage de Fontanille rencontre le même problème que plusieurs théories du cinéma : le positionnement corporel de l'énonciation se fait en privilégiant le sens de la vue qui, à travers une synecdoque dans laquelle une partie de l'expérience sensorielle

---

<sup>109</sup> Des perspectives plus « faibles » qui parlent seulement des activités cognitives générales du spectateur (BORDWELL, 1985), à des théories plus « fortes » qui placent la base de l'activité de production du sens dans le cerveau, comme le modèle proposé par Torben Grodal, du flux PECMA (GRODAL, 2009).

(qui intègre simultanément tous les « modes sensoriels ») devient le modèle pour la totalité de cette expérience. Ainsi, pour ce qui concerne le cinéma, l'auteur critique la distinction entre *focalisation* et *ocularisation* (cf. Jost), en la reformulant dans des termes de sa typologie des observateurs : le *Focalisateur* (quand le rôle d'observateur reste abstrait, sans manifestation figurative) et le *Spectateur* (quand les limites de la compétence du Focalisateur reçoivent une manifestation figurative et quand il est impliqué dans les catégories spatio-temporelles de l'énoncé). Mais, comme les perceptions de ce spectateur figurativisé sont de diverses natures, Fontanille ajoute :

*...on voit bien dans ce cas que la composante perceptive n'est rien d'autre qu'une spécification pragmatique (somatique ici) du parcours figuratif du Performateur, et que le « regard » ou « l'ouïe » ne sont que des manifestations parmi d'autres, éventuellement privilégiées par le mode d'expression, des restrictions ou augmentations de compétence pour l'observateur. On sait fort bien que les espaces pris en charge par le regard, l'ouïe, le toucher ou l'odorat sont fort différents, et qu'il faut ou bien proposer une théorie des espaces perceptifs (comme le fait G. Deleuze), ou bien adopter une stratégie qui en neutralise provisoirement les différences. Le moyen terme, consistant par exemple à privilégier le faire visuel, est une cote mal taillée. (c'est moi qui souligne) (pp. 108 – 109)*

La différence entre ces « espaces perceptifs » donne lieu à de nombreux jeux dans le cinéma (et comme on essayera de montrer, particulièrement dans le cinéma d'animation), donc il est nécessaire de reformuler le rôle de la diversité perceptive dans l'énonciation. Fontanille fait lui-même quelques remarques sur le son dans cet ouvrage. La première concerne, précisément, son idée d'un « contrat cinématographique ». Dans ce cas, l'utilisation des bruits hors diégèse sera assumée comme une rupture d'un tel contrat, la « transgression de la clause anthropomorphe », bien que cette clause soit plus faible pour les sons « diégétiques ». En effet, le sémioticien place les sons dans les trois contrats dont il vient de parler : à son avis, dans le cas du contrat filmique « le son n'a aucune part, étant donné que les savoirs perceptifs élémentaires recomposés en images grâce à ce contrat n'ont rien à voir avec les savoirs perceptifs auditifs ». Dans ce cas, « à ce niveau d'analyse, une étude acoustique, voire sémio-acoustique, montrerait comment l'observateur recompose des images auditives à partir des fréquences (...) et des amplitudes » (p. 117). Pourtant, si le contrat filmique rend compte de la perception du mouvement à partir des images fixes, le son y joue un rôle très important, particulièrement dans le cinéma d'animation, où il y a beaucoup plus de plans fixes que dans le cinéma traditionnel, et où l'impression de mouvement est donc donnée par le son.

En ce qui concerne les contrats cinématographique et narratif-discursif, Fontanille dit que le son « s'intègre sans difficulté », mais c'est parce que, pour lui, le rôle du son se limite, dans le premier, à contribuer « à disposer les figures-sources en profondeur, au même titre que la composition plastique et la construction en perspective », et dans le second à participer au montage, en assurant des transitions entre syntagmes et en soulignant des ruptures, qu'elles soient diégétiques ou pas.

Il nous faudra définir plus précisément la participation du son dans ces « contrats » (et même envisager la possibilité de concevoir d'autres contrats de type « audio-visuel », à la manière de Chion), et voir, par exemple, si la modalisation cognitive de l'espace et la typologie des effets de subjectivité sont envisageables du point de vue de la dimension sonore.

Or, on peut reformuler le rôle de la perception dans l'énonciation (et dans la sémiotique en général) sans réduire sa multiplicité à une seule instance. C'est le chemin que prend, à notre avis, Jacques Fontanille dans ses ouvrages les plus récents<sup>110</sup>. Ainsi, dans *Sémiotique du discours* l'auteur écrit :

*Dans la perspective de l'énonciation, le corps propre est traité comme un simple point, un centre de référence pour la deixis. Mais, dans la perspective des logiques du sensible, par exemple, il sera traité comme une enveloppe, sensible aux sollicitations et aux contacts venus soit de l'extérieur (sensations) soit de l'intérieur (émotions et affects). En d'autres occasions, il sera considéré comme une chair sensible et motrice, dont la plasticité dynamique lui permet de s'ajuster aux morphologies sensibles du monde naturel, ou de les adapter à lui. Dès lors, la réflexion sur la « fonction sémiotique » débouche sur une véritable sémiotique du corps signifiant. Et le corps propre n'est plus dans ce cas un simple dénominateur commun (le terme « neutre » dans la paire « extéroceptif / intéroceptif »), mais un opérateur sémiotique complexe, dont les multiples facettes (point-référence, enveloppe-mémoire, chair-mouvement, etc.) ont des fonctions bien spécifiées. (FONTANILLE, 2003 : p. 36)*

---

<sup>110</sup> Bien que, dans son livre *Sémiotique du visible*, l'auteur a commencé déjà à considérer la perception comme une dimension fondamentale dans l'activité énonciative : « si on se place du côté du discours, les phénomènes d'ondulations, de modulations et d'enchevêtrements témoignent d'une activité perceptive-énonciative qui se fait un chemin parmi les figures de la manifestation ; ces phénomènes discursifs relèvent de la *tensivité*, dont l'instance est le sujet de la *perception* » (FONTANILLE, 1995 : p. 6).



Il s'agit donc d'un projet qui sera achevé dans l'ouvrage *Corps et sens*, où Fontanille place définitivement le corps à la base de la sémiotique : « nous formons donc le projet suivant : en prolongement des éclairages théoriques qu'elle est supposée nous offrir, notamment quant au rôle du corps dans l'énonciation et la fonction sémiotique, la sémiotique du corps participera directement à la constitution d'une *syntaxe figurale*. En effet, par l'intermédiaire de la sensation et de la perception, le corps, placé au cœur de la production du sens, est susceptible de nous fournir des modèles de la schématisation, de la transformation et de la mise en séquence des figures » (FONTANILLE, 2011 : pp. 53 – 54).

Fontanille propose ici un parcours sensoriel qui place les différents « sens », non plus comme des « canaux » et mécanismes biophysiques de la sensorialité, mais comme des « champs sensoriels » du corps. Il s'agit, donc, d'une dimension « polysensorielle » de la signification et de l'énonciation, dans laquelle le corps prend position et, selon le type d'expérience, affronte les différentes propriétés des champs sensibles.

On peut replacer ici la question de l'audiovision afin de reformuler le rôle du son dans ce type d'expérience décrit par Chion de façon éclectique, en lui donnant un statut sémiotique et, surtout, une description qui dépassent l'« oculocentrisme » des théories filmiques traditionnelles. En effet, la théorie des champs sensoriels relativise l'importance que l'on a donnée à la vision en général :

*...il paraît clair qu'aucune des propriétés du champ sémiotique polysensoriel n'appartient exclusivement à tel ou tel mode sensoriel ; elle ne lui appartient que tendanciellement, en tant que manifestation dominante, et par la force d'un usage. On ne s'étonnera donc pas de constater que la vision partage bien des propriétés avec les autres sens, et qu'elle ne nous apprend plus grand-chose [une fois que les autres types de champs sont décrits] sur le champ sensible des sémiotiques-objets (p. 70).*

Par contre, pour envisager l'audiovision de cette perspective, il est nécessaire d'approfondir la description du champ « réversible et simultané » de l'ouïe, et d'aborder avec plus d'attention ses rapports avec le champ du visible. Car, bien que cette description souligne la dimension spatiale de l'écoute et des phénomènes sonores (dimension qui est restée cachée dans nombreuses théories) la relation entre spatialité et temporalité reste encore à développer.

On peut s'appuyer ici sur la phénoménologie des sons, tel comme la propose Don Ihde, qui établit également une différenciation entre les champs sensoriels. En effet, les



correspondances entre la phénoménologie de l'écoute d'Ihde et la théorie des champs sensoriels de Fontanille sont nombreuses, bien que chaque auteur fasse porter l'essentiel de sa réflexion sur des aspects différents. Par exemple, Ihde décrit lui aussi le champ sensoriel de l'ouïe comme une sphère (IDHE, 2007 : pp. 75 – 76), mais une sphère instable, sans homogénéité constante<sup>111</sup>, mais en ajoutant la « directionnalité » des sons qui touchent cette sphère. Ainsi, la forme du champ sonore est-elle non seulement enveloppante, mais aussi directionnelle.

D'après Ihde, ces propriétés de l'écoute permettent de percevoir des aspects des objets normalement attribués à la vue : on perçoit aussi des formes, des textures et des « intérieurs » des objets dans la dimension sonore. Mais surtout, l'expérience de l'écoute implique une *temporalité*, qui permet de reformuler les relations entre les champs du visible et de l'audible. Auparavant, l'auteur a défini ces activités sensorielles d'après le concept phénoménologique d'« horizon » : « L'invisible c'est l'horizon du regard. Une recherche sur l'audible est aussi une recherche sur l'invisible. Entendre rend l'invisible présent d'une façon semblable à la présence du muet dans la vision » (IDHE, 2007 : p. 51) [Traduction de l'auteur]<sup>112</sup>. Le silence est donc l'horizon de l'audible.

Ihde présente ces deux « régions » grâce aux diagrammes suivants :

---

<sup>111</sup> Dans *Corps et sens* Fontanille a décrit le champ de l'ouïe en termes similaires, mais avec la figure de la *bulle*, au lieu d'une sphère : « le champ auditif est une bulle : il généralise le principe de l'enveloppe extéroceptive, mais en même temps, il la dissocie du corps propre, en ce sens qu'il ne s'agit plus d'une enveloppe corporelle plus ou moins distante, mais d'une bulle déformable à volonté et indépendante du corps percevant » (FONTANILLE, 2011 : p. 68).

<sup>112</sup> « *The invisible is the horizon of sight. An inquiry into the auditory is also an inquiry into the invisible. Listening makes the invisible present in a way similar to the presence of the mute in vision* ».

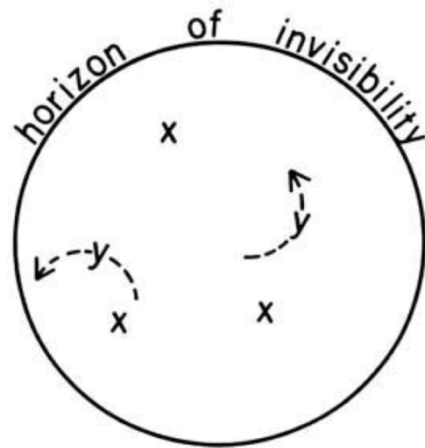


Figure 7 : L'horizon d'invisibilité

Source : [IHDE, 2007, p. 52]

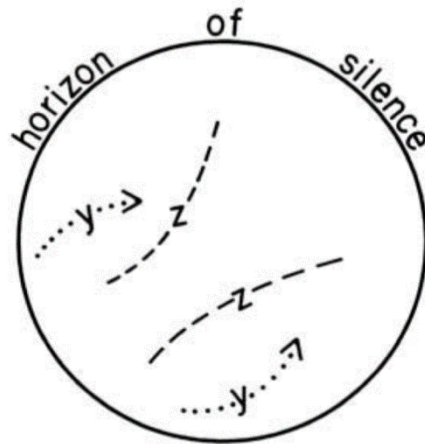


Figure 8 : L'horizon du silence

Source : [IHDE, 2007, p. 52]

Ainsi, il y a des phénomènes visibles dans le champ du regard, représentés par les X et les Y dans le cercle, dans le premier diagramme. Les X sont les entités stables, tandis que les Y sont des entités en mouvement, qui peuvent être accompagnées de sons (mais qu'on peut toujours regarder sans les écouter). Dans le deuxième diagramme, il y a aussi des Y, qui représentent les sons qui accompagnent les objets ou êtres en mouvement, mais il y a des sons sans aucune présence visuelle (les Zs). Et Ihde ajoute : « dans la mesure où tous les sons sont aussi des « événements », tous les sons doivent être considérés « en mouvement » (p. 53)<sup>113</sup>.

<sup>113</sup> « Insofar as all sounds are also “events”, all the sounds are, within the first approximation, likely to be considered as “moving” ».

Enfin, l'auteur propose une superposition des deux champs, dans laquelle il y a des objets muets (X) « fermés » à l'expérience de l'écoute, et des sons invisibles dans le champ de l'audible qui sont présents pour l'ouïe, mais absents pour l'œil. Il y a aussi des « présences » qui sont synthétisées ou présentes dans les deux champs ou régions (les Y) :

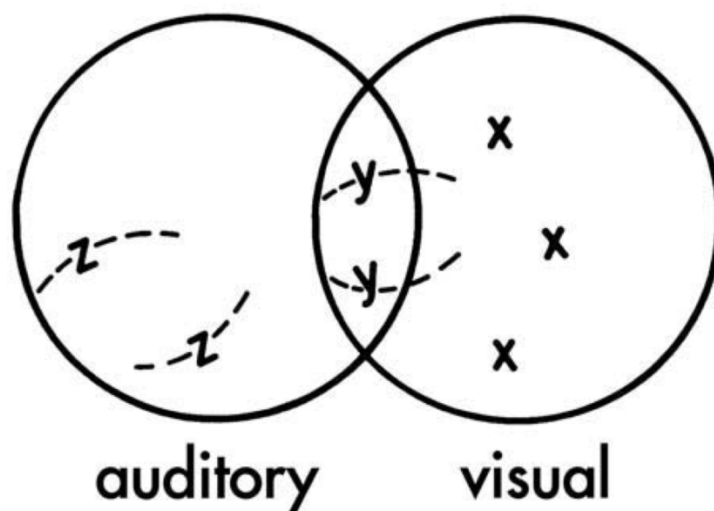


Figure 9 : Superposition de l'audible et du visible

Source : [IHDE, 2007, p. 53]

Comme l'auteur le souligne, il est nécessaire de ne pas oublier le caractère dynamique et temporel de ces relations, surtout dans la dimension sonore, dans la mesure où « le son incarne le sens du temps ».

En effet, pour Ihde l'horizon de l'audible est une *limite temporelle*. L'auteur souligne le caractère évènementiel des sons, dont découle la nécessité de les envisager dans une perspective dynamique. Pour rendre compte de la dimension temporelle des sons, Ihde suit le modèle proposé par Husserl :

*Une phénoménologie de l'expérience temporelle rencontre la notion de période temporelle (temporal span) ou durée des sons expérimentés dans l'écoute. Je n'entends pas un instant suivi par un autre ; j'entends une gestalt durable dans laquelle les modulations de la mélodie, le discours et les bruits se présentent ensemble. L'instant comme un atome de temps est une abstraction qui est liée à l'illusion d'une chose en soi. En termes de champ perceptif nous avons noté qu'une chose se produit toujours comme située dans une plus grande unité du champ. Alors, temporellement, l'instant est ici perçu comme se produisant juste dans la durée plus large d'une période temporelle, un présent vivant et étendu. Par ailleurs, selon l'analyse préalable de Husserl, cette période temporelle s'affiche comme structurée en fonction de*

*l'apparition des traits qui viennent à la perception, la protention, et du passage progressif des traits qui sortent de la présence, la rétention. À l'intérieur de la période temporelle, l'expérience continue d'une gestalt est vécue comme une succession dans l'espace d'une durée* (IDHE, 2007 : p. 89)<sup>114</sup>.

Néanmoins, cette version de la temporalité des expériences sonores doit être rapportée avec l'aspect attentionnel de l'intentionnalité de l'écoute, qu'Ihde appelle le « foyer temporel » (temporal focus). Comme activité intentionnelle, l'écoute peut se fixer sur certains aspects de l'expérience. Par exemple, normalement on perçoit le langage oral comme un flux –sans distinguer chaque phonème ou syllabe individuellement–, mais on peut aussi se concentrer pour isoler des unités sonores. Dans ce cas, le flux devient « arrière-plan », tandis que le phonème spécifique passe au foyer temporel.

Ce « *temporal focus* » peut-être étroit, fin ou ample, selon l'intentionnalité de l'auditeur. Quand le foyer est très étroit, l'arrière-plan devient plus marginal. Ainsi la notion d'*instant* se rapporte à la durée comme le premier plan se rapporte à l'arrière-plan.

En outre, Ihde souligne qu'il y a aussi une *forme* du foyer temporel qui le donne sa direction. Ici Ihde remarque qu'il est nécessaire de ne pas prendre la *linéarité* comme la seule caractéristique des sons : « Dans la temporalité auditive, la période temporelle se révèle comme contenant une multiplicité d'événements auditifs qui sont intentionnellement classés. Il est à la fois une simultanéité et une succession » (IHDE, 2007, p. 111). Dans ce contexte, Ihde se demande comme placer le foyer temporel dans la période : est-ce qu'il est au centre,

---

<sup>114</sup> « A phenomenology of experienced temporality soon comes on the notion of a *temporal span* or duration of sounding that is experienced in the listening. I do not hear one instant followed by another; I hear an enduring gestalt within which the modulations of the melody, the speech, the noises present themselves. The instant as an atom of time is an abstraction which is related to the illusion of a thing in itself. In terms of a perceptual field we have noted that a thing always occurs as situated within a larger unity of a field; so temporally the use of instant here is perceived to occur only within the larger duration of a temporal span, a living present. Moreover, according to Husserl's prior analysis (to be modified below), this temporal span displays itself as structured according to the onset of features coming into perception, protention, and the phasing and passing off of features fading out of presence, retention. Within the temporal span the continuing experience of a gestalt is experienced as a succession within the span of duration ».

comme le foyer de la perception visuelle, ou peut-il se déplacer à travers la période temporelle ?

Chez Husserl le champ temporel est caractérisé comme limité par des horizons, et le « maintenant » est canalisé vers le futur par un ensemble de *protentions* ou *attentes* qui appartient à notre expérience temporelle. Dans cette perspective, les protentions sont des « intentions temporelles vides », qui cherchent à accéder à la présence, mais qui peuvent en cela être accomplies ou frustrées. Les protentions sont des structurations attentionnelles orientées vers le futur.

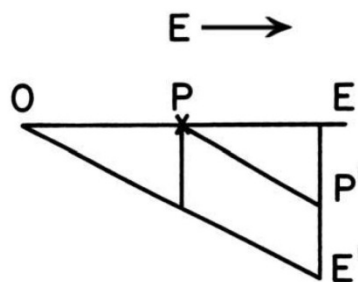


Figure 10 : Diagramme du temps chez Husserl

Source : [IHDE, 2007, p. 92]

Dans le diagramme de la figure 3 la ligne OE représente les « points-maintenant » qui constituent le foyer temporel. La Ligne OE' c'est la profondeur du foyer, qui contient à l'intérieur des réverbérations du présent que vient de passer, des phénomènes qui coulent mais qui sont maintenus jusqu'à leur disparition dans la mémoire. La flèche montre que, chez Husserl, « la 'direction' du temps est du futur vers le passé ». En d'autres termes, le futur est ce qu'on passe, en le transformant en présent et après en passé.

Selon Ihde, il est possible de faire la mise au point étroitement bien dans le point d'apparition du phénomène, ou bien dans son extinction. Cela donne lieu au deuxième diagramme proposé par Ihde, où F est le foyer temporel, a est le point-source de l'événement sonore et b est son point d'extinction :

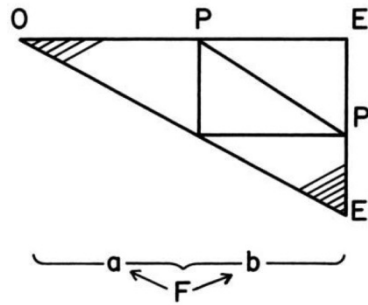


Figure 11 : Diagramme du foyer temporel

Source : [IHDE, 2007, p. 93]

À partir de cette variation du schéma de Husserl, Don Ihde propose une autre modification, pour inclure les « horizons » des deux directions du foyer temporel : l'horizon du « pas encore (vide) », où on est face au futur ouvert, et l'horizon du « pas plus », où on est face au passé dont les rétentions trouvent leur limite dans la mémoire (les rétentions ne sont pas des souvenirs, quand elles passent cette limite elles deviennent souvenirs, ou tombent dans l'oubli). Le diagramme de la figure 5 représente ce champ entre horizons.

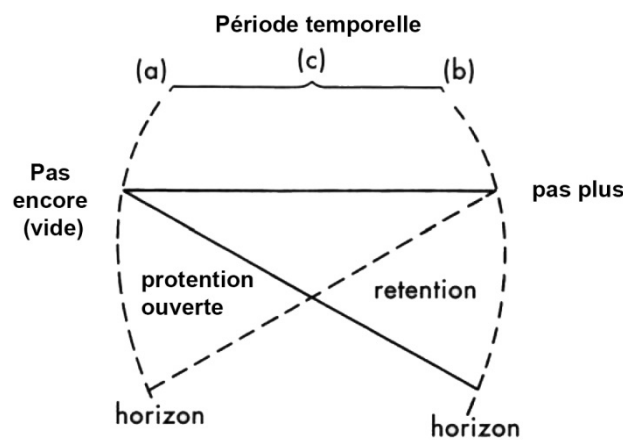


Figure 12 : Diagramme de la période temporelle

Source : [IHDE, 2007, p. 95]

Ce que l'auteur veut souligner avec ces diagrammes c'est que, à la différence du foyer visuel qui est « centré », le foyer temporel de l'ouïe ne l'est pas, puisqu'il peut se concentrer sur les extrêmes de la période temporelle en question, ou même il peut être ouvert à tout le champ, dans un phénomène particulier de l'écoute qu'Ihde appelle l'« état de champ » (*field state*). C'est le cas de certaines expériences musicales, dans lesquelles l'attention de l'auditeur est pleine et ouverte.

Il n'est pas gratuit que les exemples préférés par les phénoménologues à l'heure de parler du temps soient des phénomènes sonores. La temporalité complexe de l'écoute permet la simultanéité, mais aussi la linéarité, et cette double direction de l'attention sonore est la base l'aspectualité, comme on va voir dans le chapitre 3. On va voir aussi que la compréhension et manipulation des caractéristiques du champ sonore sont une partie très importante du dispositif audiovisuel. Par exemple, le pouvoir d'anticipation de la musique (basé sur l'adressage du foyer temporel vers l'horizon des protentions) est un phénomène fortement étudié (HURON, 2006). Pour l'instant je veux souligner que les propositions d'Ihde complètent clairement celles de Fontanille. Les phénomènes ici décrits sans tout le détail qu'ils méritent laissent pressentir le besoin d'une approche tensive. En outre, on y trouve confirmation de la « porosité » catégorielle entre les différents types de champs du sensible, qui échangent et modulent les mêmes propriétés : notamment, selon Ihde, les propriétés aspecto-temporelles, qui reprennent en les reconfigurant des propriétés temporelles (comme la rémanence ou l'anticipation) identifiées par Fontanille dans le champ olfactif. Ici on veut surtout souligner l'intérêt de cette approche pour décrire l'expérience « audiovisuelle ». Mais pour comprendre une telle expérience, où les relations entre les images et les sons sont construites et « modulées » selon des intentions expressives, il est nécessaire de considérer au moins deux aspects entrelacés, mais différents : d'un côté, il s'agit d'interpréter la forme dont le « texte filmique », comme on l'entend dans la perspective greimassienne, « simule » l'expérience du spectateur. Et de l'autre, il faut rendre compte d'une *situation d'exhibition* dans laquelle une *expérience est activée et modulée*, et une pratique est déployée. L'audiovision nous place, donc, entre différents niveaux de pertinence tels que définis par Jacques Fontanille dans *Pratiques sémiotiques* (FONTANILLE, 2008) : dans le trajet qui va du texte filmique à la pratique de l'audiovision.

### **L'énonciation audiovisuelle comme pratique(s) sémiotique(s)**

On a déjà fait remarquer que le problème des théories de l'énonciation cinématographique consiste à privilégier un énonciataire réel, tout en se référant également à des instances énonciatrices entendues comme « simulacres ». Ce conflit se traduit, en termes sémiotiques, dans la confusion entre l'« énonciation énoncée » et le contexte de communication effective dans lequel l'énonciation se produit :

*Au cours de ces vingt dernières années, dans le cadre de la sémiotique post-greimassienne, la conception de l'énonciation a évolué ; ces évolutions sont allées de pair avec une*

*reconsidération des rapports entre énoncé et contexte de communication. Le contexte de communication, qui avait toujours été considéré comme une question résolue à l'intérieur de la sémantique de l'énoncé, est désormais devenu un niveau pertinent de l'analyse.* (ANGENOT, DONDERO, JOACHIM, & SHIRKHODAEI, 2013 : p. 536)

C'est donc une nouvelle perspective, celle de l'« énonciation en acte », ce qui a permis développer une théorie où nous pouvons situer les rapports entre texte (filmique) et corps percevant. Cette perspective est présentée dans *Pratiques sémiotiques*, où Jacques Fontanille expose sa proposition des niveaux de pertinence :

*La signification « en acte » imputable par l'analyse à un texte ne peut être rigoureusement observée et saisie, de fait et de droit, qu'au niveau des pratiques, et non des textes-énoncés proprement dit. Au niveau de pertinence des textes l'« en acte » ne relève au mieux que de la substance, au pire de la spéculation animiste. C'est donc l'expérience de l'« en acte » (de l'activité vivante et vécue) qui donnera lieu, par schématisation, au plan d'immanence des « scènes pratiques ». (FONTANILLE, 2008 : p. 26)*

Fontanille propose donc une *hiérarchie des plans d'immanence* qui permet d'identifier les niveaux de pertinence pour l'analyse sémiotique, qui résultent de sa proposition de commencer toute réflexion sur l'origine du sens dans l'expérience sensible, dans une perspective inspirée par la phénoménologie :

*... partir de l'« apparaître » des phénomènes qui s'offrent aux divers modes de la saisie sensible, c'est déjà définir un plan originel pour le plan de l'expression : on admet ici en quelque sorte que le plan de l'expression présuppose une expérience sémiotique, et la solution qui pourrait en découler consisterait alors à s'interroger sur les niveaux de pertinence de cette expérience, en se demandant sous quelles conditions ils peuvent être convertis en « plans d'immanence » pour l'analyse sémiotique ». (FONTANILLE, 2008 : pp. 17-18)*

C'est dans ce contexte qu'on peut situer l'énonciation dans sa dimension « vivante », comme une énonciation « en acte » :

*Au niveau de pertinence du texte, l'énonciation n'est pertinente que si elle y est représentée (énonciation énoncée), alors que l'énonciation dite « présupposée » est un pur artefact sans observables. Mais au niveau de pertinence des objets-supports, voire des pratiques qui les intègrent, l'énonciation retrouve toute sa pertinence : les acteurs y retrouvent un corps et une identité, l'espace et le temps de l'énonciation leur procurent un ancrage déictique, et les actes*



*mêmes de l'énonciation peuvent s'inscrire figurativement dans la forme d'expression issue de la morphologie matérielle même des objets d'inscription.* (FONTANILLE, 2008, p. 47)

L'énonciation donc se situe entre les niveaux de l'objet-support et de la pratique. Mais il est nécessaire de noter que d'une part, dans le cas du cinéma le niveau de l'objet-support est constitué par une hétérogénéité d'éléments : en ce qu'on a appelé l'expérience de l'audiovision dans la projection dite « traditionnelle »<sup>115</sup>, au moins : l'écran, les hautparleurs, la salle elle-même et les conditions perceptives (la position frontale de l'écran, l'obscurité, le son distribué ou concentré, etc.).

D'autre part, on ne peut pas oublier les instances actanciennes de l'énonciation : l'instance énonciatrice et l'instance énonciataire. Dans la perspective des pratiques sémiotiques, on peut décrire ces deux faces comme participant à des pratiques concurrentes. Comme nous l'avons déjà montré, les éléments relatifs à la production peuvent être abordés au début sous la perspective du texte-énoncé, pour décrire l'« énonciation énoncée » avec ses différentes instances énonciatrices (narrateurs, monstateurs, bruiteurs, etc.)<sup>116</sup>. On peut aussi identifier des simulacres textuels des énonciataires : il y en a un « corps virtuel » comme le corps décrit par Casetti, pure « textualité », mais construit aussi à partir de l'expérience sensible. Ou, peut-être mieux, comme la prothèse symbolique que propose Bettetini.

Mais, en suivant Bettetini, cette prothèse doit être actualisée par des corps « réels », circonscrits dans la scène pratique de la « projection », où le texte filmique prend la forme de son support hétérogène : la lumière sur l'écran et les sons provenant des hautparleurs sont gérés dans un espace de projection tridimensionnel où les sons et les images rencontrent un

---

<sup>115</sup> Pour l'instant, nous allons simplement prendre l'exemple du cinéma traditionnel. Plus tard, lorsque nous allons introduire le concept de dispositif, nous pourrions nous occuper des différentes situations d'énonciation audiovisuelle.

<sup>116</sup> Cette perspective permet aussi d'aborder sémiotiquement l'énonciation dite « présupposée », comme l'a fait François Jost dans son œuvre *Sous le cinéma, la communication* où, dans le chapitre quatre, il propose de se placer « du côté de la production, de la fabrication du film » pour décrire « le regard de l'auteur. Pas le regard de l'auteur sur ses personnages non plus, mais le regard sur ses acteurs, sur la réalité qu'il filme » (JOST, 2014, p. 89). Effectivement, la production des films peut être décrite dans le niveau de pertinence des pratiques. On reviendra sur ce sujet dans le chapitre prochain.

corps « tout percevant »<sup>117</sup> ; et il faut enfin prendre en considération l'horizon stratégique de cette scène pratique, l'horizon où elle interfère avec toutes les autres scènes pratiques concurrentes et/ou concomitantes réunies dans l'expression « aller au cinéma ».

L'avantage d'aborder l'activité spectatorielle de cette façon c'est qu'elle permet montrer que, comme le dit Aumont, « le spectateur de film ne va pas au cinéma seulement pour comprendre, mais pour mettre en jeu sa sensibilité, c'est-à-dire en fin de compte, son corps »<sup>118</sup>. Mais le défi est grand parce que, comme l'a déjà dit Fontanille, « l'introduction du « sensible » et du « corps » dans la problématique sémiotique entraîne en effet quelques difficultés qui n'ont pas été résolues jusqu'à présent, et qui tiennent au fait que ce « sensible » et ce « corps » ne sont pas nécessairement représentés dans le texte ou dans l'image pour être « pertinents », notamment quand il s'agit d'articuler l'énonciation sur une expérience sensible et sur une corporéité profonde. Il ne suffit pas, par exemple, de renvoyer les notions relevant de la « phorie » et de la « tensivité » à une couche « proto-sémiotique » pour leur procurer un statut clair et opératoire » (FONTANILLE, 2008, p. 48).

Néanmoins, il y a ici une clé pour aborder la scène sémiotique du dispositif audiovisuel. Selon Fontanille, « ...au niveau supérieur, celui des pratiques sémiotiques (les pratiques de « production de sens », les pratiques interprétatives, notamment), les valences sont celles de la perception qui, à l'intérieur d'une pratique, permettent de construire, de saisir ou d'imaginer l'univers du texte dont la pratique fait usage ; les valences trouvent alors toute leur pertinence : un univers sensible est donné à appréhender [...] et c'est alors que les valences jouent leur rôle, comme « filtre » praxique de la construction axiologique » (FONTANILLE, 2008, p. 48).

La hiérarchie des plans d'immanence permet donc placer les différents éléments de l'expérience audiovisuelle dans ses niveaux de pertinence correspondants, mais aussi de contempler ses relations dans une coexistence syntagmatique des sémiotiques-objets. Dans la

---

<sup>117</sup> Pour reprendre une expression de Christian Metz, « un écran et un espace de projection sensorielle et imaginaire à la fois, où se construit l' « autre scène », celle, quasi-fantasmatique, où s'associent toutes les sensations, les affects et les cognitions engendrées dans la scène primaire » (METZ, 1977b).

<sup>118</sup> (AUMONT, 2011 : p. 134). Aumont fait ce commentaire à propos de l'ouvrage de Raymond Bellour, *Le corps du cinéma. Hypnoses, émotions, animalités* (BELLOUR, 2009) où ce dernier associe aussi le concept de dispositif cinématographique à l'hypnose.

perspective de Fontanille, cette coexistence est résolue par syncrétisme, grâce à la construction des parcours génératifs des plans d'expression et de contenu, qui permet convertir « un ensemble hétérogène en 'ensemble signifiant' : ainsi, successivement, le texte, l'image, l'objet, la scène pratique, la stratégie d'ensemble, sont traités comme des ensembles signifiants, dont on peut proposer une description actancielle, modale, passionnelle, figurative et énonciative, quel que soit le niveau de pertinence où on se place » (FONTANILLE, 2008, p. 51).

Une stratégie similaire est proposée pour deux autres questions en rapport avec l'hétérogénéité : la polysensorialité et les hétérogénéités énonciatives. Examinons en particulier cette dernière, par sa relation étroite avec notre sujet. On a déjà vu que l'énonciation audiovisuelle, dans toutes ses dimensions, est hétérogène : des différentes instances productrices jusqu'aux instances des pratiques concurrentes dans la scène pratique de la réception. Pour rendre compte de cette multiplicité on trouve ici une solution proche de celle de Christian Metz dans son dernier ouvrage : « pour restituer toute la complexité de ces voix multiples, fût-elle cacophonique, il faut supposer *un espace d'énonciation impersonnel et collectif*, au sein duquel des stratégies d'ajustement, de conflit et de collusion se font jour et donnent forme, au niveau supérieur, à une 'représentation partagée' » (FONTANILLE, 2008, p. 55). Et plus bas, il ajoute : « La multiplicité de voix énonciatives subit donc le même sort que la polysensorialité et que la multimodalité : l'hétérogénéité est résolue (et non réduite) par redistribution à l'intérieur d'un plan d'immanence qui reconfigure la diversité en distribution cohérente » (FONTANILLE, 2008, p. 56)<sup>119</sup>.

Effectivement, plusieurs auteurs ont parlé de polyphonie dans l'énonciation cinématographique. Parmi d'autres, François Jost a développé la théorie de l'énonciation polyphonique de Ducrot, en l'étendant au niveau de l'image (JOST, 1992, pp. 78 - 89). Dans un sens similaire, Jean Châteauvert a proposé une théorie de l'énonciation polyphonique pour penser la voix « over » (CHATEAUVERT, 1996). Et on a déjà parlé de la proposition d'Alain Boillat, et ces différents « niveaux énonciatifs » (BOILLAT, 2006, pp. 99 - 110). Maintenant, on peut revenir sur cette dernière, pour mettre en discussion quelques problèmes particuliers de l'énonciation audiovisuelle.

---

<sup>119</sup> Nous soulignons.

Dans sa théorie, Boillat partage partialement la défense de l'impersonnalité de l'énonciation filmique et sa critique à l'égard de l'anthropomorphisme impliqué dans des notions comme « grand imagier » ou « méga-narrateur » : « l'analyse gagne en effet à évincer ces connotations, non pas, comme le prétend Metz, parce que le cinéma repose sur des moyens techniques (car chacun d'eux est soumis à une utilisation humaine), mais parce qu'elles tendent à nous éloigner du marquage énonciatif tel qu'il se présente dans le texte filmique même à des niveaux d'élaboration qui ont peu à faire avec l'instance désignée ». Boillat choisit de se « limiter à l'analyse concrète des marques textuelles plutôt que tenter de définir l'hypothétique représentation qu'un spectateur (lui-même dans ce cas tout aussi hypothétique) se fait de l'énonciateur » (BOILLAT, 2006, pp. 98-99). C'est dans ce contexte qu'il parle des niveaux énonciatifs, associés avec ces marques : « Je propose, pour déterminer ces niveaux, de les calquer sur quatre éléments constitutifs du discours filmique : l'iconique, le verbal (mentions graphiques et orales), le musical et le bruitage. Cette division quadripartite offre une base relativement palpable qui facilite la prise en compte de la technique même. À chacune de ces composantes correspond donc un type de marquage énonciatif ». (BOILLAT, 2006, p. 100)<sup>120</sup>.

On trouve ici un problème difficile à encadrer sous les théories traditionnelles de l'énonciation cinématographique, mais aussi sous la proposition de Fontanille : où peut-on placer ces « marques énonciatives » ? Dans le texte filmique, comme le dit Boillat ? Mais, où se trouve la dimension sonore d'un tel texte ? Dans les haut-parleurs ? Mais font-ils partie de l'objet-support ? Ou dans le corps du spectateur qui écoute (en faisant donc partie de la « pratique ») ? Comment identifier la frontière entre ces niveaux ? L'hétérogénéité énonciative, donc, n'est pas si facile à réduire, comme il n'est pas toujours facile de distinguer la frontière entre les niveaux de pertinence.

---

<sup>120</sup> Il y a un intérêt particulier dans la proposition de Boillat : il prend en compte une dimension énonciative qui nous concerne précisément : « J'insiste sur le cas du bruitage, trop souvent omis dans la théorie: le bruit n'est bien sûr pas totalement inféodé à l'image (sinon quelle serait la fonction des bruiteurs ?), mais constitue une composante à part entière de l'énoncé filmique. Cet aspect, couramment masqué par des tentatives de naturalisation, nous apparaît parfois lorsqu'un cinéaste joue sur celui-ci, comme le fait souvent Jacques Tati à des fins comiques, ou Kurosawa dans *Dodes Kaden* (1970) lorsqu'on entend momentanément le bruit d'un tram qui n'existe que dans l'imagination du personnage ».

Or, il y a un concept dans l'histoire des théories filmiques qui peut nous aider pour affronter ces problèmes. Il s'agit d'un concept qui permet de lier la question technique et matérielle de l'expérience audiovisuelle avec sa dimension énonciative. Il s'agit du concept de *dispositif*.

## **Le dispositif et l'énonciation audiovisuelle**

*C'est donc d'abord au niveau de l'appareillage que le cinéma fonctionne comme un langage ; inscription d'éléments discontinus dont l'effacement dans le rapport qui s'institue entre eux serait producteur de sens. (BAUDRY, 1970, p. 20)*

La théorie du dispositif a ses origines dans la période de confluence des théories sémiologiques et de la psychanalyse, dans les années soixante-dix. Le théoricien le plus cité à cet égard est Jean-Louis Baudry, qui a donné au concept de dispositif le sens que nous discutons ici, dans deux articles fondateurs (BAUDRY, 1970), et (BAUDRY, 1975). L'autre auteur qui a contribué à développer ce concept est Christian Metz, dans les articles recueillis dans *Le signifiant imaginaire* (METZ, 1977). La « tournure psychanalytique » de la sémiologie du cinéma participe à sa manière à ce qu'on a appelé le parti pris pour le spectateur<sup>121</sup>, à savoir l'intérêt prioritaire pour l'activité perceptive et psychologique des êtres humains dans la salle de cinéma<sup>122</sup>.

---

<sup>121</sup> Qui sera continué par les théories cognitives à partir des années quatre-vingt-dix.

<sup>122</sup> Ici il est nécessaire de faire une remarque. On peut dire qu'il y a un sens restreint et un sens large du concept de *dispositif*. Nous optons ici pour le sens restreint, celui de Baudry et Metz, et que dans le contexte anglo-saxon ce qui est connu comme la « apparatus theory », développé à partir des auteurs français par Teresa de Laurentis, Stephen Heath et les collaborateurs de son ouvrage (DE LAURENTIS & HEATH, 1980). C'est aussi l'orientation du numéro 1 de 2003, du dossier de *Cinémas : revue d'études cinématographiques*, consacré à la théorie des dispositifs cinématographiques, dont on va citer plusieurs articles. Ce sens restreint a été étendu au champ des médias en général. En ce qui concerne le sens large, le concept du dispositif vient de la philosophie, et a une place central chez Michel Foucault (DELEUZE, 2003, pp. 316-425) ; Giorgio Agamben a fait des rapports entre ce concept et la philosophie du technique de Martin Heidegger (AGAMBEN, 2014). On peut trouver aussi le sens étendu chez Jean-François Lyotard par rapport à la psychanalyse et l'esthétique (LYOTARD, 1994), ou chez Althusser et son concept d'appareils idéologiques d'état (ALTHUSSER, 1976, págs. 67-125). En général ce sens élargi a à voir avec les dispositions, agencements et configurations institutionnelles, spatiales, politiques ou psychologiques des formes

Dans son texte fondateur (1970), Baudry inclut dans son modèle du dispositif cinématographique tant la production (avec un accent particulier sur le travail de la caméra) que la réception (projection) : « D'une façon générale, nous distinguons l'*appareil de base*, qui concerne l'ensemble de l'appareillage et des opérations nécessaires à la production d'un film et à sa projection, du *dispositif*, qui concerne uniquement la projection et dans lequel le sujet à qui s'adresse la projection est inclus » (BAUDRY, 1975, pp. 58 - 59). En ce qui concerne la production, l'auteur prend comme point de départ les rapports entre les techniques de représentation de l'espace que le cinéma hérite de la peinture et la construction d'un type particulier de sujet<sup>123</sup>. Mais la particularité du cinéma réside dans la projection : « ...ce qui était déjà à l'œuvre comme fondement constitutif de l'image perspectiviste, c'est-à-dire l'œil, le « sujet » est relancé, libéré (comme une réaction chimique libère une substance) par l'opération qui transforme des images successives, discontinues (en tant qu'images isolées, elles n'ont à proprement parler pas de sens, pas d'unité de sens pour le moins) en continuité, mouvement, sens. La continuité rétablie, c'est à la fois sens et conscience qui le sont » (BAUDRY, 1978 : pp. 19 – 20).

---

sociales, qui configurent la subjectivité et les comportements des individus dans un contexte historique et social. Bien sûr, la limite entre les deux sens n'est pas toujours nette, dans la mesure où les technologies des médias et des spectacles font partie des configurations incluses dans le sens large du *dispositif*, comme on peut le voir dans le dossier de la revue *Hermès* n° 25, *Le dispositif entre usage et concept* où on trouve des applications et développements du concept dans la sociologie, les études de médias et la technologie (JACQUINOT-DELAUNAY & MONNOYER, 1999) ; (PEETERS & CHARLIER, 1999). Bien que ce ne soit pas l'objectif de cette thèse, je pense que le sens large du concept a aussi un intérêt sémiotique, en particulier pour une sémiotique des pratiques, des stratégies et des formes de vie (Fontanille, *Formes de vie*, 2015).

<sup>123</sup> L'intérêt de Baudry est de souligner la dimension idéologique de la technique. C'est ainsi qu'il affirme : « ...le tableau de chevalet en présentant un ensemble immobile et sans intervalle, élabore une vision pleine qui répond à la conception idéaliste de la plénitude et de l'homogénéité de l'« être » et, pour ainsi dire, il est le représentant de celle-ci » (BAUDRY, 1978 : p. 17). En ce sens son approche est plus proche des positions « culturalistes » de la perspective. Pour une lecture sémiotique de ces positions, voir (FONTANILLE, 1989), en particulier la deuxième partie, consacrée à la théorie de l'observateur dans la peinture : « L'argument de l'exactitude mathématique porte en fait sur la *forme de l'expression*, et plus particulièrement sur une figure technique, alors que l'argument de la relativité socioculturelle porte à la fois sur la *forme du contenu* dans ses relations avec les figures plastiques, et sur les *langages connotatifs* associés à la représentation de l'espace » (p. 71).

En plus, selon Baudry « les mouvements de la caméra semblent-ils réaliser les conditions les plus favorables à la manifestation d'un sujet transcendantal. Il y a à la fois fantasmatisation d'un réel objectif qui, réduisant ses pouvoirs de contrainte, semble augmenter d'autant les possibilités ou la puissance du sujet » (BAUDRY, 1978 : p. 20). Cette idée est développée par Metz dans *Le signifiant imaginaire* : « Le spectateur, en somme, s'identifie à lui-même, à lui-même comme pur acte de perception (comme éveil, comme alerte) : comme condition de possibilité du perçu et donc comme à une sorte de sujet transcendantal, antérieur à tout *il y a* » (METZ, 1977 : p. 69)<sup>124</sup>.

On peut se souvenir de l'interprétation psychanalytique du dispositif de projection qui a rendu Baudry célèbre, en associant le projecteur, la salle obscure et l'écran à la caverne de Platon<sup>125</sup>, et simultanément à la « phase du miroir » proposée par Lacan. Toutefois, comme l'a bien dit Jacques Aumont, à propos de cette conception du dispositif « on ne saurait donc trop insister sur le caractère spéculatif de ces propositions, que ne corrobore aucune expérimentation de type scientifique, et qui dressent du dispositif cinématographique et de son sujet spectateur un tableau excessivement centré sur l'asservissement de l'un par l'autre, même dans des versions plus soucieuses de rester concrètes » (AUMONT, 2011 : p. 132).

Dans le contexte anglo-saxon, il y a eu des auteurs qui ont validé et développé le concept de dispositif ou d'appareil, mais il y a aussi de nombreuses critiques, pointées surtout sur les filiations psychanalytiques, ou sur le manque de perspective historique. Parmi les premières, on trouve des auteurs qui ont proposé la version américaine de la théorie du dispositif, sous le nom de « apparatus theory ». Il y a même un ouvrage avec ce titre : le recueil édité par Teresa de Laurentis et Stephen Heath en 1980. Ce recueil inclut des articles des auteurs les plus importants de la théorie du cinéma à l'époque, y compris des théoriciens plus proches des théories cognitives qui prendront la relève quelques années plus tard, comme Joseph et Barbara Anderson, les principaux représentants de l'approche écologique de la perception appliquée au cinéma (voir ANDERSON, 1996). Il y a aussi des théoriciens de l'importance de Peter Wollen, Dudley Andrew, pour le contexte américain, et Jean-Louis Comolli (COMOLLI, 1971/2009), pour le contexte français.

---

<sup>124</sup> Metz développe son approche de l'énonciation filmique dans son dernier ouvrage, "L'énonciation impersonnelle" (Metz, 1991). Il n'y a pas accent particulier sur le dispositif, mais sur la façon dont le film "parle de lui-même", en particulier à travers une dimension méta-discursive.

<sup>125</sup> Reprise et développée dans son second article (Baudry, 1975).



Dans ce recueil je voudrais attirer l'attention sur trois articles, pour leur relation avec ma recherche : le premier est celui de Douglas Gomery, *Towards an Economic History of the Cinema: The Coming of Sound to Hollywood* (GOMERY, 1980). Dans cet article, Gomery propose une lecture historique et économique du développement des techniques de son synchronisé à Hollywood, avec l'intérêt d'aller plus loin que le *Chanteur du Jazz* (1927), le film ou la majorité des historiens (y compris, Comolli) placent l'origine du parlant. Gomery souligne le rôle des courts métrages de Studios comme Warner Bross ou Fox, qui ont expérimenté différentes techniques de synchronisation. Gomery remarque le rôle de l'économie dans la compréhension de l'évolution des dispositifs. Aussi son emphase sur la dimension historique le conduit à contester quelques critiques faites à la théorie du dispositif qu'on va mentionner ensuite.

Le deuxième c'est l'article de Mary Ann Doane, *Ideology and the Practice of Sound Editing and Mixing* (DOANE, 1980), où l'auteure poursuit le chemin de Baudry, Comolli et d'autres auteurs de cette théorie, en démasquant l'« idéologie bourgeoise » qui soutient le discours des monteurs du son dans l'industrie hollywoodienne. Et bien que les arguments soient presque les mêmes (la recherche d'une unité métaphysique, l'impérialisme du visible, le masquage de l'artifice et des conventions pour exhiber une fausse « naturalité », etc.), ce qui est intéressant dans le texte de Doane est l'insistance sur l'aspect matériel et technique de la pratique de montage sonore, qu'elle connaît bien et qu'elle connecte avec des questions conceptuelles de la création audiovisuelle.

Le troisième article que je veux souligner est celui de Kristin Thompson, *Implications of the Cel Animation Technique* (THOMPSON, 1980), tout simplement parce que c'est l'un des rares textes qui prennent en compte l'animation du point de vue d'une théorie des dispositifs. Bien qu'elle ne s'arrête pas sur la dimension sonore des cartoons, sa conclusion peut être extrapolée à toutes les techniques impliquées dans la production des dessins animés :

*...les dessins animés utilisent certains appareils qui sont potentiellement très perturbateurs (par exemple, des mélanges de systèmes de perspective, les indices de vitesse antinaturalistes). Comme on peut espérer, à l'intérieur du système hollywoodien classique, cependant, les motivations narratives et comiques adoucissent ces perturbations. Même un film si radical dans ses techniques comme Duck Amuck reste assez lisible pour un public habitué à regarder Daffy dans ses films plus caractéristiques. Comme toujours, les techniques et la technologie*



*de films ne sont pas radicales par elles-mêmes; elles le deviennent lorsqu'elles sont utilisées dans la structure d'un film complet. (THOMPSON, 1980, p.118)<sup>126</sup>*

Du côté des critiques on trouve l'approche dite cognitiviste, dont l'ouvrage inaugural est le livre édité par David Bordwell et Noël Carroll, *Post-theory. Reconstructing Film Studies* (BORDWELL & CARROLL, 1996). Consacré spécifiquement à la critique de la théorie du dispositif c'est l'article de Vance Kepley Jr., *Whose Apparatus ? Problems of Film Exhibition and History* (le dispositif de qui? quelques problèmes de la projection des films et l'histoire) (KEPLEY Jr., 1996). Dans cet article, Kepley conteste la possibilité d'extrapoler le modèle du dispositif donné par Baudry :

*Comment et pourquoi les gens ont choisi se rassembler en présence d'un projecteur de cinéma au fil des ans sont des questions que des historiens sont encore à explorer. La théorie du dispositif a esquivé ces bases dans le parcours d'appropriation d'un type de projection et lui a attribué le statut d'un archétype. On peut argumenter qu'il y a une distinction utile entre le « spectateur théorique » postulé par la psychanalyse (et développé par Baudry) et le « spectateur empirique », qui doit être en quelque sorte confinée au domaine des historiens. La nature anhistorique de la théorie du dispositif est ainsi pensée pour être justifiable en tant qu'elle crée un modèle qui n'est pas limité à une période historique particulière. Le cas échéant, cependant, la théorie du dispositif est une simple extrapolation d'une situation historique limitée. Elle prend le type de projection la plus familière pour son auteur (et, probablement, pour son lectorat immédiat), le type pratiqué dans les théâtres de première diffusion dans l'Europe de l'après-guerre, par exemple, et elle l'a attribué à la totalité du cinéma. Quoi qu'il en soit, la théorie du dispositif a été à son insu un produit de l'histoire et elle est donc aussi limitée dans ses possibles applications que n'importe quelle approche empirique dans la littérature sur le sujet. (KEPLEY Jr., 1996, p. 546)<sup>127</sup>*

---

<sup>126</sup> “We have seen how cartoons use some devices which are potentially very disruptive (for example, mixtures of perspective systems, anti-naturalistic speed cues). As we might expect within the classical Hollywood system, however, narrative and comic motivations smooth over these disruptions. Even a film as radical in its devices as *Duck Amuck* remains quite readable to an audience accustomed to watching *Daffy* in his more characteristic films. As always, film techniques and technology are not in themselves radical; they become so only when used within the structure of a complete film”.

<sup>127</sup> “How and why people have chosen to gather in the presence of a movie projector over the years are issues historians are still exploring. Apparatus theory skipped over such basics in the course of appropriating one type of exhibition and assigning it the status of an archetype. It might be argued

Cette citation met en évidence plusieurs problèmes dans la théorie de Baudry. Kepley donne comme alternative ce qui dans le contexte anglo-saxon s'appelle les Études d'audiences (*audience studies*), en particulier dans sa ligne historique.

Pourtant, la théorie du dispositif n'a pas à voir seulement avec des publics –bien que la « scène pratique » de la séance de cinéma fasse partie de ses intérêts-. L'aspect qui nous intéresse plus dans l'approche de Brophy est la relation qu'il établit entre les composants techniques du dispositif et *la production d'une expérience de sens*. Il en résulte que les critiques adressées au concept ne sont pas insurmontables. Jacques Aumont lui-même reprend ce concept pour parler des images, et pas seulement au cinéma, mais en le séparant de ses résonances psychanalytiques pour en faire un modèle plus général applicable à des situations différentes.

Les précisions apportées par cet auteur sur le dispositif sont d'un grand intérêt pour nous. Aumont avance quatre traits qui permettent de définir tous types de dispositifs médiatiques et artistiques :

1. **Un dispositif a toujours deux dimensions**, l'une **physique** (des appareils, des matériaux, des lieux, des moments), l'autre **mentale** (des pensées, des sensations, des émotions, des affects). Plus précisément, le dispositif est ce qui sert à mettre en

---

*that there is a useful distinction to be maintained between the "theoretical spectator" postulated by psychoanalysis (and developed by Baudry) and the "empirical spectator", which is somehow understood to be the separate domain of historians. The ahistorical nature of apparatus theory is thus thought to be justifiable since it creates a model which is not limited to particular historical periods. If anything, however, apparatus theory simply extrapolated from a historically limited situation. It took the kind of exhibition most familiar to its author (and presumably to its immediate readership), the kind practiced in first-run theaters in postwar Europe, for example, and it assigned that model to the whole of cinema. If anything, apparatus theory was unwittingly a product of history and thus is as limited in its useful applications as any of the empirical accounts in the literature".* Kepley remplace cette idée de dispositif avec des alternatives théoriques sociologiques et historiques sur les douanes de projection, la constitution des publics du cinéma, et le rôle des salles et théâtres dans contextes historiques et géographiques concrets. Ces dimensions, bien que très intéressants peuvent être inclus, dans les niveaux des « stratégies » où des « formes de vie », selon le cas. Voir (FONTANILLE, 2015).

relation ces deux dimensions, en faisant se rencontrer un certain agencement matériel et un agencement de formes de pensées ou d'affects –sans qu'on puisse pour autant dire que l'un de ces agencements est entièrement déterminé par l'autre.

2. **Un dispositif implique une organisation de l'espace.** Les appareils qu'il mobilise, et les formes de pensée qu'il gère sont pris dans une véritable structure. Inversement, l'existence même d'un dispositif voué à un certain effet structure l'espace (/temps) où il s'inscrit.

3. Il ne s'agit pas, cependant, d'une structure statique, qui se contenterait d'encadrer une pratique. Un dispositif est plutôt comparable à un moteur : il **implique un dynamisme, la mise en jeu d'une énergie motrice.**

4. **Il vise un résultat** (et généralement l'obtient) : un dispositif est destiné à placer un sujet humain dans des *dispositions* (de corps et d'esprit) qui l'inclinent à penser à certains objets, à engager certaines actions ou à éprouver certaines passions (AUMONT, 2011 : pp. 123-124) (c'est l'auteur qui souligne).

Le point numéro 1 nous permet de concevoir le dispositif comme une véritable fonction sémiotique avec un plan de l'expression (la dimension physique, avec des appareils, des matériaux, des lieux, des moments), et un plan du contenu (la dimension mentale des pensées, sensations, émotions, et affects). Les points 2 à 4 vont dans le sens de l'inscription corporelle de la sémiosis et de l'énonciation : la structuration de l'espace et du temps, le dynamisme et l'énergie motrice et la dispositionalité corporelle et mentale des sujets humains qui participent constituent les traits d'une scène où les corps vivants interagissent et se constituent en un foyer de sens.

La question est, maintenant : où placer le dispositif cinématographique dans la hiérarchie proposée par Fontanille? Il ne s'agit pas de placer le dispositif comme un niveau de pertinence indépendant, comme a essayé de faire Semir Badir avec les *médias*, bien que le cinéma soit souvent défini comme un *médium*. Fontanille a déjà montré que, bien que « ce choix présente quelques avantages en matière de description (...) il n'est pas sûr que les

médias, qui comprennent à la fois les « canaux », les « formats » et des règles de la pratique, c'est-à-dire des propriétés hétérogènes, puissent faire l'objet d'une analyse continue au sens strict » (FONTANILLE, 2008 : p. 23).

Nous proposons de comprendre le dispositif comme *la scène sémiotique de la pratique* de réception cinématographique, une scène et une pratique qui, comme l'a dit Frank Kessler, ne sont pas « intemporelles », mais, comme toutes les pratiques, sont historiquement situées :

*On voit dès lors que la notion de dispositif, conçue comme l'interrelation entre une technologie, un mode de présentation, un mode d'adresse, une forme filmique et un positionnement de l'instance spectateur, demande finalement à être « historicisée ». En abandonnant l'idée d'une prédisposition psychique fondamentale, inscrite pour ainsi dire ontologiquement dans le média en tant que tel et régissant la cinématographie dans son ensemble, on pourrait par ailleurs distinguer au cours de l'histoire du cinéma plusieurs dispositifs, chacun étant rendu spécifique par son historicité.* (KESSLER, 2003 : p. 25)<sup>128</sup>

On a déjà vu que, pour Baudry, l'appareil de base comprend la production et la réception, tandis que le dispositif se concentre sur la scène de réception, y compris le spectateur. Plus récemment Jean-Pierre Sirois-Trahan a proposé de distinguer les notions de « dispositif de réception » et de « dispositif de production »<sup>129</sup>. Dans une perspective diachronique, cet auteur croit que le concept de dispositif permet de déduire la réception du cinéma : « si l'hypothèse que le dispositif est 'une métonymie et une métaphore du spectateur' est juste, qu'il est une métaphore en quelque sorte du fonctionnement mental du spectateur en plus d'en être une synecdoque (le spectateur est une partie du dispositif), pourquoi ne pas la

---

<sup>128</sup> En suivant cette idée de Kessler, Sirois-Trahan identifie cinq types de dispositifs qui existaient dans le cinéma des premiers temps : dispositifs de représentation scénique ; dispositifs panoramiques et véhiculaires ; dispositifs mixtes théâtre/cinéma ; dispositifs de vues sonores synchronisées ; et dispositifs identificatoires (les plus proches de la description de Baudry). Ces derniers étaient les seuls qui ont triomphé et qui ont donné lieu au cinéma institutionnel tel qu'il est connu aujourd'hui. Les autres dispositifs, que Sirois-Trahan dénomme « mode de représentation en trompe-l'œil », « survivent aujourd'hui dans les expositions universelles, les parcs d'attractions, certaines nouvelles technologies (casque virtuel, simulateur, IMAX, etc.) et au Futuroscope de Poitiers » (Sirois-Trahan, 2003 : pp. 169 – 170).

<sup>129</sup> Cet auteur parle même d'un *dispositif de distribution*, bien qu'il ne le décrive pas (Sirois-Trahan, 2003, p. 163).

prendre au pied de la lettre ? De telle sorte que chaque variation du dispositif soit un indice de la variation de la réception spectatorielle » (SIROIS-TRAHAN, 2003, p. 161).

Si cette hypothèse est correcte, nous pouvons revenir sur la question de l'énonciation audiovisuelle à partir d'une dimension qui nous permet d'intégrer les questions textuelles proprement dites avec les problèmes liés à la scène pratique de la réception audiovisuelle. Dans la perspective des pratiques sémiotiques, on peut décrire ces deux faces comme des pratiques concurrentes. Comme nous l'avons déjà montré, les éléments relatifs à la production peuvent être abordés sous la perspective du texte-énoncé, pour décrire l' « énonciation énoncée » avec ses différentes instances (narrateurs, monstateurs, bruiteurs, etc.), en relation avec les simulacres de(s) énonciataire(s)<sup>130</sup>. Il y a ici un « corps virtuel » comme le corps décrit par Casetti, pure « textualité » (mais construit aussi à partir de l'expérience sensible). Ou, peut-être mieux, comme la prothèse symbolique qui propose Bettetini. D'autre part, les éléments relatifs à l'exhibition relèvent des corps « réels », circonscrits dans la scène pratique de la « projection » (dans le cas du cinéma « en salle »).

Il y a ici, on l'a déjà dit, une fonction sémiotique, dont le plan du contenu est constitué d'affects, de synesthésies, d'effets corporels, de sensations, d'émotions, de pensées ; c'est-à-dire, entre autres, les produits de ce que Roger Odin appelle la *mise en phase* (ODIN, 2000, pp. 37 - 46). En ce sens, la fonction sémiotique du dispositif peut être rapprochée de celle des sémiotiques plastiques.

Dans la perspective d'une sémiotique des pratiques, la scène pratique qui se met en place comprend, donc : un texte filmique ; un corps « tout percevant » ; et il faut enfin prendre en considération l'horizon stratégique de cette scène pratique, l'horizon où elle interfère avec toutes les autres scènes pratiques concurrentes et/ou concomitantes : un horizon de projection dans une salle de cinéma et, en général, toutes les dimensions dont parle Kepley, et qu'on peut connecter avec les niveaux des stratégies et des formes de vie<sup>131</sup>.

---

<sup>130</sup> Mais maintenant il est possible, dans la perspective des pratiques, de faire aussi des analyses des pratiques de production effectives.

<sup>131</sup> Dans *Langage et cinéma*, Christian Metz déjà faisait la distinction entre le fait cinématographique et fait filmique : « le film n'est qu'une petite partie du cinéma, car ce dernier figure un vaste ensemble de faits dont certains interviennent avant le film (infrastructure économique de la production, studios, financement bancaire ou autre, législations nationales, sociologie des milieux de décision, état

Mais pour démontrer la pertinence et l'effectivité du concept de dispositif, il faut montrer comment un tel concept permet d'aborder la totalité des pratiques sémiotiques liées à l'expérience audiovisuelle, mais aussi d'intégrer les différents niveaux de pertinence dans lesquels on peut rendre compte suffisamment de cette expérience dans une perspective sémiotique. À cette fin, dans le chapitre suivant, nous proposons l'analyse d'un cas particulier, que nous permettra de confronter l'intérêt de la proposition de Fontanille autour des niveaux de pertinence du plan de l'expression, mais aussi d'identifier les problèmes de cette proposition par rapport au cinéma en général, et au cinéma d'animation en particulier.

Pour l'instant, nous voulons proposer quelques lignes dans lesquelles on peut développer le concept du dispositif audiovisuel dans un cadre sémiotique.

Premièrement –et pour conclure partiellement la discussion sur l'énonciation audiovisuelle– le dispositif dit « de production », a à voir avec l'énonciation « en acte ». Greimas et Courtés ont déjà défini la production comme « un tout, et qui, situé dans l'instance de l'énonciation, abouti à la formation de l'énoncé (phrase ou discours) » (GREIMAS & COURTÉS, 1993). Mais, dans le cas du cinéma, on a déjà souligné la multiplicité d'instances d'énonciation qui produisent des énoncés dits « syncrétiques ». Chaque instance peut-être décrite en soi –et non pas seulement celles qui ont été privilégiées par la tradition de la sémiologie filmique, comme le travail de la caméra ou le montage, mais aussi, par exemple, les différentes instances du travail avec le son–, pour, après, les intégrer dans un parcours collectif d'énonciation que donne lieu aux films achevés. Ici, on peut identifier les différentes

---

technologique des appareils et des émulsions, biographie des cinéastes, etc.), d'autres après le film (influence sociale, politique, idéologique du film sur les différents publics, « patterns » de comportement ou de sentiment induits par la vision des films, réactions des spectateurs, enquêtes d'audience, mythologie des stars, etc.), d'autres enfin pendant le film, mais à côté et en dehors de lui : rituel social de la séance de cinéma (moins lourd que dans le théâtre classique, mais qui tient de cette sobriété même son statut dans la quotidienneté socio-culturelle), équipement des salles, modalités techniques du travail du projectionniste, rôle de l'ouvreuse (c'est-à-dire sa fonction dans divers mécanismes économiques ou symboliques, que n'entamerait pas son inutilité pratique), etc. » (METZ, 1971, p. 7). À son époque, l'option évidente était pour le *discours signifiant localisable*, le film comme *texte*, objet sémiotique d'une recherche sémiologique. Aujourd'hui, nous pouvons inclure les dimensions « cinématographiques » dans une recherche sémiotique, en prenant soin de les placer dans les niveaux de pertinence correspondants.

pratiques sémiotiques de la production audiovisuelle et les rapports –solidarité, complémentarité, confrontation, etc. – entre elles<sup>132</sup>.

De la même manière, une réflexion sémiotique sur le dispositif permet aborder d'une nouvelle façon la question des différentes catégorisations des types de discours audiovisuels. Il y a eu, pour mettre un exemple pertinent pour notre recherche, de grandes difficultés pour définir l'animation comme une forme audiovisuelle. Finalement, le seul consensus est que l'animation se distingue par ses modes de production, c'est-à-dire, par les pratiques créatives particulières qui donnent le mouvement à des objets –dessins, modèles, figures de papier, etc.- inanimés. Certaines de ces pratiques sont partagées avec le cinéma en général. D'autres appartiennent exclusivement au cinéma d'animation. En prenant les pratiques de production du cinéma d'animation comme sémiotique-objet, on peut tester la pertinence de cette approche dans le domaine des théories filmiques.

Mais le concept du dispositif permet aussi d'examiner le rôle des matières et objets-supports dans la production audiovisuelle. Comme on va essayer de montrer avec notre première analyse, dans quelques types d'animation il s'agit de travailler et d'expérimenter avec les supports traditionnels du cinéma : la pellicule, avec ses deux bandes d'enregistrement. Cela a permis de trouver de nouvelles formes de relation entre l'image et le son, exclusives –depuis les débuts, jusqu'à l'arrivée de la révolution numérique– du dispositif d'animation directe<sup>133</sup>.

Le concept du dispositif est fondamental pour articuler les différents niveaux de pertinence : comme scène pratique, on peut décrire les relations entre les actants énonciateurs, leur faire productif, les supports et matériaux qui sont configurés par ce faire et les propriétés des faires et des supports dans les formes audiovisuelles du texte filmique qui en résultent.

---

<sup>132</sup> Il est toujours possible de parler de l'énonciation audiovisuelle dans la perspective des « auteurs », même d'après la théorie du dispositif. C'est le cas du livre *Perception-cinéma: Les enjeux stylistiques d'un dispositif*, où Clélia Zernik essaie de montrer comment les dispositifs conduisent la perception des spectateurs en différentes directions (ZERNIK, 2010). Pourtant cette auteure donne aussi une certaine importance à la dimension collective, culturelle, de l'énonciation, notamment quand elle analyse les films des cinéastes orientaux.

<sup>133</sup> Dans le prochain chapitre on va parler du dispositif « animatique ».



Quant au dispositif de réception (du côté des énonciataires) ce concept permet d'aborder plusieurs questions auparavant exclues par la sémiologie du cinéma, mais remises en scène par d'autres théories. Pour commencer, il est nécessaire de trouver des alternatives théoriques à la psychanalyse de Baudry pour tenir compte des émotions et de la corporalité réelle des spectateurs dans la scène sémiotique de la réception audiovisuelle<sup>134</sup>. Ici, la sémiotique des pratiques peut combler le hiatus entre les théories fondées sur les techniques audiovisuelles et les théories psychologiques et cognitives sur l'expérience des spectateurs. On peut voir sur ce sujet le travail de Laurent Jullier, avec un fort accent sur la dimension sonore, notamment (JULLIER, 1995) ; (JULLIER, 2002, pp. 51-59 ; 75-86) ; et (JULLIER, , 2012, pp. 172-320). Dans le contexte anglo-saxon un ouvrage récent très intéressant sur l'« esthétique occulte » de la synchronisation son/image au cinéma consacre un chapitre aux rapports entre bande-son et bande-image d'un point de vue cognitif (DONNELLY, 2014, pp. 14-43). Le recueil *Lowering the Boom. Critical Studies in Film Sound* inclut aussi des articles qui prennent en considération des aspects rapportés au dispositif, bien que ce ne soit pas l'approche dominante dans le livre (BECK & GRAJEDA, 2008). La contribution de Beck au volume comporte des remarques sur système Dolby Stereo (BECK, 2008); l'article de James Lastra réfléchit sur la pratique du *Sound Design* (LASTRA, 2008) ; et, surtout, la contribution de Paul Grainge sur la publicité des entreprises du système Dolby et ses publicités (GRAINGE, 2008).

Un autre axe de recherche apparaît si on décide de prendre en compte les propositions de Vance Kepley Jr. citées ci-dessus. Tout le domaine des « audiences studies » est donc ouvert à la recherche sémiotique, ainsi que les études historiques des dispositifs et des publics. Sur l'exhibition des films et les séances du cinéma, on peut voir, par exemple (RAE HARK, 2001), ou (WALLER, 2001) ; sur l'expérience des spectateurs, voir (PLANTINGA, 2009), bien que ces études fussent consacrées au contexte américain. Or si le dispositif-cinéma traditionnel n'est qu'un parmi d'autres, il faut explorer ces autres configurations de sens, ces autres pratiques d'audiovision tout au long de l'histoire. Sur une révision historique des

---

<sup>134</sup> Cet aspect de la théorie de Baudry a été le plus critiqué par les cognitivistes. Voir, notamment, l'attaque de Noël Carroll sur les relations établies par Baudry entre le cinéma et le rêve (CARROLL, 1988). Pourtant il y a aussi des auteurs qui aujourd'hui défendent la pertinence de la théorie de Baudry. Par exemple, dans un livre consacré spécifiquement aux transformations du son multipistes au cinéma après la révolution numérique, Mark Kerins utilise la théorie du dispositif à la manière de Baudry, et répond pour une part aux objections de Carroll à Baudry (KERINS, 2011, pp. 277-296).



dispositifs audiovisuels qui ne se limite pas seulement au cinéma, le travail de Zielinski est de grand intérêt (ZIELINSKI, 1999).

Dans les chapitres suivants, nous allons essayer d'aborder certaines de ces questions à partir d'analyses concrètes. Le premier cas est le plus proche de ces pages d'introduction théorique. Avec l'analyse d'un dispositif audiovisuel concret, nous allons tester la pertinence de ce concept, aussi bien que l'intérêt d'appliquer la théorie des pratiques et des niveaux de pertinence dans le domaine de la sémiotique audiovisuelle. Ainsi, la recherche créative de l'animation expérimentale nous va permettre d'identifier un type particulier de dispositif de grand intérêt pour mieux comprendre ce qu'on a appelé l'audiovision.

## Chapitre 4. Dispositifs audiovisuels et cinéma d'animation

Comme on a essayé de le montrer dans le chapitre précédent, le concept du dispositif doit être diversifié au moins en deux directions : la première, selon la critique la plus répandue, pour identifier plus qu'un seul dispositif, puisque la description classique de Baudry est restreinte à une unique espèce parmi d'autres. De nombreux auteurs ont étudié la diversité des dispositifs dans le cinéma des premiers temps<sup>135</sup>, mais cette description peut être faite tant en diachronie qu'en synchronie. Ainsi, on peut étudier l'évolution (et multiplication) des dispositifs à travers l'histoire, ou travailler sur des dispositifs à l'heure aujourd'hui.

La deuxième dérive de la distinction proposée par Jean-Pierre Sirois-Trahan entre dispositifs de production et dispositifs de réception. D'après cet auteur,

*Le dispositif de production comprend tout ce qui sert à produire un film : la caméra certes, mais plus largement l'ensemble du studio de production, et toutes les opérations qui s'y effectuent. Donc un ensemble matériel et de processus non mentaux qui forment le dispositif de production matériel. Ce dispositif de production a également un biais mental dans la mesure où les producteurs (soit le metteur en scène, le cameraman, le producteur, les acteurs, etc.) ont non seulement une image de leurs spectateurs, mais également des formes, des règles symboliques, des conventions et des contraintes que ces derniers vont mettre en branle, ou subir, pour produire du sens. (SIROIS-TRAHAN, 2003, p. 163)*

D'autre part, le dispositif de réception, selon la distinction de Sirois-Trahan, est présenté à partir de la difficulté de parler de la réception dans un sens historique :

*L'absence des spectateurs rend obligatoire, selon moi, la description des dispositifs pour rendre compte de la réception (...) Du reste, à moins de faire des recherches sur l'activité cognitive ou de recourir à des enquêtes sociologiques, démarches impossibles pour le passé,*

---

<sup>135</sup> On a déjà mentionné le travail de Kessler sur les dispositifs des premiers temps (Kessler, 2003). À part lui, il y a une grande tradition de réflexions sur les dispositifs et pratiques de la première époque du cinéma. Voir par exemple (GUNNING, 2006) or (GAUDREAU, 2008). Sur la composante sonore spécifiquement, voir (ABEL & ALTMAN, 2001) ; (ALTMAN, 2007) ; (BARNIER, 2011). Finalement, sur le son au cinéma d'animation des premiers temps, voir (CURTIS, 1992) et (GOLDMARK, 2007).

*la réception spectatorielle ne peut qu'être un déduit ou un construit. Il faut aborder la réception indirectement.* (SIROIS-TRAHAN, 2003, p. 161)

On a déjà vu son idée du dispositif de réception selon laquelle les variations du dispositif sont des indices de la variation de la réception spectatorielle. Néanmoins, il y a un décalage entre les dispositifs de production et de réception que le cinéma dit « classique » ou institutionnel a caché, puis ce dispositif, le plus étudié par Baudry et autres théoriciens, implique selon Sirois-Trahan une coïncidence presque parfaite entre les dispositifs de production et de réception : « On pourrait ainsi définir le cinéma institutionnel comme la coïncidence réglée des dispositifs mentaux de la production et de la réception sur le modèle narratif » (SIROIS-TRAHAN, 2003, p. 163).

On trouve chez Sirois-Trahan une autre distinction importante pour regarder les différences entre les dispositifs audiovisuels. Il s'agit d'une distinction qu'on a déjà trouvée chez Aumont : la dimension physique et la dimension mentale du dispositif. Sirois-Trahan parle de « dispositif mental » et de « dispositif matériel » (SIROIS-TRAHAN, 2003, p. 163). On peut donc proposer la table combinatoire suivante (en prenant comme exemple le dispositif du cinéma traditionnel) :

Tableau 1: Les dimensions du dispositif du cinéma traditionnel

	Dispositif de production	Dispositif de réception
Dispositif matériel	Appareils et processus (pratiques) : caméra, microphones, lampes, etc. Tournage, montage, etc.	Appareils et processus (pratiques) : salle (obscure), projecteur, écran, hautparleurs, chaises, etc. Silence, observation collective, etc.
Dispositif mental	Conventions et usages de réalisation : le « langage narratif cinématographique ».	Conventions et usages de réception : le « langage narratif cinématographique ».

En termes sémiotiques, le dispositif de production a à voir avec un ensemble de pratiques coordonnées par des stratégies. Ces pratiques ont pour objet la transformation d'un support (la pellicule) en un texte audiovisuel, à travers des appareils et techniques de manipulation des images et des sons. Dans le cinéma PVR<sup>136</sup>, la manipulation de l'image se

<sup>136</sup> Dans son œuvre *Le cinéma d'Animation*, Sébastien Denis utilise l'expression de production en *Prises de Vues Réelles* (PVR), que nous utilisons aussi (DENIS, 2011).

déroule à travers la photographie. Dans une grande partie du cinéma d'animation, la photographie est aussi le *médium* de production visuel principal, mais avec une variation technique importante : la capture des photogrammes se fait image par image. En ce qui concerne le son, il y a un enregistrement des phénomènes sonores par des microphones, qui est aussi inscrit sur la pellicule. Dans le cinéma traditionnel dit « analogique », du côté de la production il y a donc un objet-support constitué par une bande-image et une bande-son, travaillées par des pratiques d'enregistrement avec la médiation des appareils qui automatisent les procès. Mais tandis que dans le cinéma PVR l'enregistrement des images et des sons peut-être simultané –et avoir la même source– dans l'animation, la relation entre images et sons, du point de vue de la production, est une complète indépendance. Ces différences se trouvent principalement dans le dispositif matériel de production. Du côté du dispositif matériel de réception ces différences ont tendance à passer inaperçues.

Dans un certain sens, le dispositif mental dont parle Sirois-Trahan peut être entendu comme la *praxis énonciative* du cinéma<sup>137</sup> : la dimension de « la signification comme fait culturel » (GREIMAS & FONTANILLE, 1994, p. 75) est conçue, à l'intérieur du dispositif, comme étant déterminée également par les usages et par les technologies. Il s'agit d'un cercle où l'introduction d'une invention technique donne lieu à des possibilités expressives qui, à leur tour, peuvent conduire à la limite l'usage d'une technologie, en exigeant des inventions, dans le dynamisme continu de l'histoire des dispositifs. Mais dans chaque dispositif la relation entre le dispositif mental de production, et celui de la réception n'est pas toujours symétrique. Chaque dispositif –et chaque contexte d'usage– gère de façon différente ces relations.

---

<sup>137</sup> Pour des définitions sémiotiques de ce concept, voir (GREIMAS & FONTANILLE, 1994 : pp. 75-77); (FONTANILLE & ZILBERBERG, 2004, pp. 161-190) ; (BERTRAND, 1993). Dans le champ de la sémiologie et des théories du cinéma, on peut trouver des équivalents de l'idée de praxis énonciative chez des auteurs comme Noël Burch dans son livre *Une praxis du cinéma* (BURCH, 1986), chez Christian Metz, depuis son texte inaugural de la sémiologie du cinéma, *Le cinéma : langue ou langage ?*(METZ, 1964) en passant par son ouvrage *Langage au cinéma* (METZ, 1971) jusqu'à son dernier livre, *L'énonciation impersonnelle ou le site du film* (METZ, 1991). Chez Metz il y a toujours eu la préoccupation pour identifier le lieu des usages et des conventions du cinéma dans le cadre de la question de la signification. Pourtant, la relation d'une telle praxis avec la médiatisation technologique inhérente au cinéma n'a été jamais entièrement résolue.

L'existence de ces différences conduit à proposer des dispositifs propres de l'animation, englobés sous une catégorie générale : un dispositif qui nous oblige même à reformuler la conception du dispositif cinématographique en général.

### **Le(s) dispositif(s) animatique(s)**

Le théoricien australien Philip Brophy a proposé le terme de « dispositif animatique » pour se référer à l'animation, en opposition au « dispositif cinématique ». Ce dernier est le dispositif traditionnel du cinéma PVR, basé sur le modèle perceptif de l'*animisme*<sup>138</sup> : l'illusion ou la mimésis du mouvement, rendue possible par un appareil d'enregistrement. Le dispositif animatique, en revanche, a à voir avec le *dynamisme* : le mouvement conçu comme une force, un travail, le résultat d'une construction (l'animation comme procès créatif).

Néanmoins, dans un sens technique tout dispositif cinématographique est un dispositif animatique, puisque cela signifie la construction artificielle du mouvement image par image. De ce point de vue, même le cinéma PVR devrait être considéré comme animation, puis il implique la mise en succession de photographies statiques à travers un artifice technique. Le dispositif animatique est donc au cœur du dispositif cinématographique. Cependant, d'un point de vue phénoménologique le dispositif de production de films est assumé par le cinéaste comme la capture de blocs de mouvement, comme l'enregistrement d'un « temps réel » qui est fixé dans le film.

Par contre, l'animateur construit le mouvement et la temporalité image par image, donc il est tout à fait conscient qu'il travaille avec un temps artificiel. Ceci est un avantage, car cela implique que l'animateur soit considéré comme un « constructeur d'action » et pas comme un simple enregistreur :

*Un dispositif animatique devrait être similairement une machine générative d'effets comme le dispositif cinématographique, mais une machine que serait intéressée par les cadres, images, coupes et parties plus comme des événements et des occurrences que comme des éléments et des composants ; syntonisée plus avec la vitesse et le tempo de la fragmentation qu'avec la séquence formelle ou l'organisation structurelle des fragments ; impliqué dans le film et la photographie plus comme une transition que comme un procès ; et axé sur l'animation plus*

---

<sup>138</sup> Ceci est un paradoxe, parce que le terme « animation » lui-même vient de l'idée de « animer l'inanimé ».

*comme une méthode de la caricature que comme une apparition de la vraisemblance.*  
(BROPHY, 1991, p. 68)<sup>139</sup>

Mais cette caractéristique du travail de l'animateur a un autre avantage, d'un grand intérêt pour notre recherche : cette prise de conscience de l'artifice visuel permet à l'animateur de mettre également l'accent sur la production sonore, qui joue un rôle clé dans la construction audiovisuelle de l'animation.

L'idée de Brophy est de promouvoir dans le spectateur l'assimilation du dispositif animatique, tout comme il a toujours assimilé le dispositif cinématographique. Dans nos termes, il s'agirait de positionner et stabiliser un *dispositif mental de réception animatique*. D'après Brophy, il est nécessaire de ne pas seulement « regarder » les films, pour prendre conscience de l'intégrité perceptuelle de la matérialité et de la textualité des objets cinématographiques. Et ce sont les conventions et les techniques du « savoir-faire » de l'animation qui nous donneront les moyens pour arrêter de simplement regarder ou voir les films. Une idée proche de Chion et de son concept d'audiovision.

Dans la mesure où le cinéaste a à traiter avec le "temps réel" de l'image-mouvement, tandis que l'animateur doit construire le mouvement en manipulant les images image par image, seul ce dernier nous permet de mettre entre parenthèses l'illusion cinématographique pour arrêter de simplement regarder les films. Et c'est dans la compréhension du rythme tel qu'il est ainsi construit pour l'animation, qu'il est plus facile de passer de la simple visualisation d'un film à la compréhension de sa temporalité.

Pour définir le dispositif animatique, Brophy souligne deux prémisses : 1) le rythme est l'expérience du temps. À travers des composants comme le tempo, le beat, l'accent, etc., on construit le sens de la temporalité. 2) Le mouvement est l'expérience de l'espace. Mais l'espace ne peut être signifiant que quand on le traverse, et cela prend du temps. La temporalité dérivée de la traversée de l'espace est effectuée par le mouvement des airs, des

---

<sup>139</sup> *“An animatic apparatus would be a similarly generative machine of effects to that of the cinematic apparatus, but one that is interested in frames, images, cuts and parts more as events and occurrences than elements or components; attuned more to the speed and tempo of fragmentation than the formal sequencing or structural organization of fragments; concerned with film and photography more as a transition than a process; and focused on animation more as a method of caricature than an apparition of lifelikeness”.*

paramètres et des dimensions par rapport au mouvement du sujet. Et cela introduit aussi le rythme dans l'espace.

Brophy propose la caméra multiplane du studio Disney comme exemple du fonctionnement du dispositif animatique. Bien que cet appareil ait été utilisé pour « affiner l'effet de réalisme » du mouvement des différents plans, Brophy souligne que cela a créé l'effet de mouvement à travers des interactions rythmiques internes plutôt que des relations visuelles externes :

*La caméra multiplane d'animation peut-être proposée comme une métaphore étrangement autoréflexive du dispositif animatique dans la mesure où cette machine d'effets génératifs met en évidence l'animation comme un modus operandi technotextuel –une prise de conscience de la façon dont la technologie et la textualité sont inséparables, de la façon dont les concepts de mesure / effet sont plus importants ici que des questions de forme / contenu. (BROPHY, 1991, p. 72)<sup>140</sup>*

Voici une idée très intéressante pour notre recherche : selon Brophy, dans le dispositif animatique le niveau de pertinence de la textualité ne peut être séparé ni du niveau de pertinence de la pratique dont il est le résultat, ni du niveau du support. Dans le cas de la caméra multiplane, le double intérêt de libérer la caméra du statisme de l'animation traditionnelle<sup>141</sup>, et de construire une « interface » entre la visualité et le rythme (tel que cela a été proposé par Oskar Fischinger) a permis une mobilisation de la dynamique grâce à laquelle l'espace et le temps deviennent des renforcements rythmiques mutuels<sup>142</sup>.

Mais le plus intéressant dans le dispositif animatique matérialisé dans les inventions du studio Disney apparaît quand on regarde la façon dont son créateur a conçu la manipulation de la bande-son, la composition musicale et les rapports narratifs entre l'image et le son, notamment en ce qui concerne la question du synchronisme :

---

<sup>140</sup> “*The multiplane animation camera can be posited as a strangely self-reflexive metaphor for the animatic apparatus in that this machine of generative effects evidences animation as a techno-textual modus operandi –a realization of how technology and textuality are inseparable, of how concepts of measure/effect are more pertinent here than issues of form/content*”.

<sup>141</sup> En suivant le besoin, selon Brophy, d'Abel Gance pour le cinéma PVR. Voir p. 72.

<sup>142</sup> Sur la camera multiplane, voir aussi (LUCENA JÚNIOR, 2002, p. 111-119).

« Le problème du synchronisme est aussi double : (a) le son et l'image doivent éventuellement s'ajuster « en synchronie » ; et (b) ni l'enregistrement visuel ni le sonore ne peuvent être conçus sans une conception de la façon dont chacun devrait être lié à l'autre » (Brophy, 1991, p. 75)<sup>143</sup>. La solution trouvée a été l'utilisation du métronome :

*Le Studio Disney a été le premier en découvrir que le métronome pouvait donner une lecture du temps avec la même pertinence pour l'animateur que pour le compositeur. Cela signifiait que, une fois que tous les deux avaient convenu dans une feuille de repère avec toutes les indications de beat appropriées pour l'action du story-board, le compositeur et l'animateur pourraient alors partir travailler séparément. L'idéal d'une relation symbiotique entre la musique et le graphisme a été initialement conçu à travers une relation industrielle entre le compositeur et l'animateur. (BROPHY, 1991, pp. 75 - 77)<sup>144</sup>*

L'idée d'un dispositif animatique –défini initialement par le dispositif de production– est valable pour la totalité du cinéma d'animation. Dans ce sens, du point de vue du dispositif animatique mental, il n'y a pas une différence profonde entre la technique appelée « *mickey mousing* » des films de Disney, et les recherches expérimentales du cinéma d'animation dit « abstrait » dans lesquelles il y a une synchronisation parfaite entre l'image et le son, bien que dans le premier cas le recours technique qui permet la synchronisation est le métronome, tandis que dans le second le recours technique est la construction sonore directement sur la pellicule. Ainsi il est possible de postuler l'existence de divers dispositifs animatiques, différenciés selon les quatre critères indiqués dans la table combinatoire qu'on a déjà formulée, et en fonction du niveau de pertinence qu'on choisit pour faire l'analyse.

En revenant sur la pratique appelée « *mickey mousing* » je voudrais souligner le fait qu'elle est un cas très intéressant d'une invention qui est assimilée comme convention

---

<sup>143</sup> “The problem of synchronism also was twofold: (a) sound and image had to eventually match up ‘in-synch’; and (b) neither visual nor sound recording could be engineered without a concept of how each should relate to the other”.

<sup>144</sup> “The Disney Studio was the first to realize that the metronome could give a reading of time which had equal relevance to the animator and the composer, which subsequently meant that once both had agreed on a cue sheet with all appropriate beat indications for the storyboard action, the composer and the animator could then go off and work separately. The ideal of a symbiotic relationship between the musical and the graphic was initially realized through an industrial relationship between composer and animator”.



expressive tant dans le dispositif de production que dans le dispositif de réception. Du côté de la production, le procès était complexe et impliquait, en plus du métronome, l'utilisation de feuilles à barres (« *bar sheets* ») créées par Wilfred Jackson pendant l'animation de *Steamboat Willie* et standardisées ensuite pour toutes les animations du studio, notamment les « *Silly Symphonies* », caractérisées par le synchronisme musique/image : (JACOBS, 2015, p. 59). Dans la feuille à barres « Chaque carré représente une mesure, mais sans la notation musicale ». Jacobs ajoute une citation d'un journal des ingénieurs du cinéma de l'année 1933, ou William Garity explique la technique :

*Le directeur et le musicien travaillent à la main, mesure par mesure, image par image; chacun essayant de régler son problème particulier pour répondre aux exigences de l'histoire. Lorsque la mise en page de la feuille est terminée, le directeur a son image complètement aménagée sur le cadre, et le musicien a sa partition à la note près. De légères modifications peuvent être apportées plus tard, pour tenir compte des exigences qui peuvent survenir lorsque les images sont animées.* (JACOBS, 2015, p. 59)<sup>145</sup>

Ici je reproduis l'image d'une feuille à barres correspondant au court métrage *Santa's Workshop* (L'atelier du père Noël) inclus par Jacobs dans son étude du phénomène (Figure 6) :

---

<sup>145</sup> “*The director and the musician work hand in hand, measure by measure, frame by frame; each one trying to adjust his particular problem to meet the demands of the story. When the layout sheet is completed, the director has his picture completely laid out to the frame, and the musician his master score to the note. Slight changes may later be made in order to accommodate the exigencies that may arise when the pictures are animated*”.

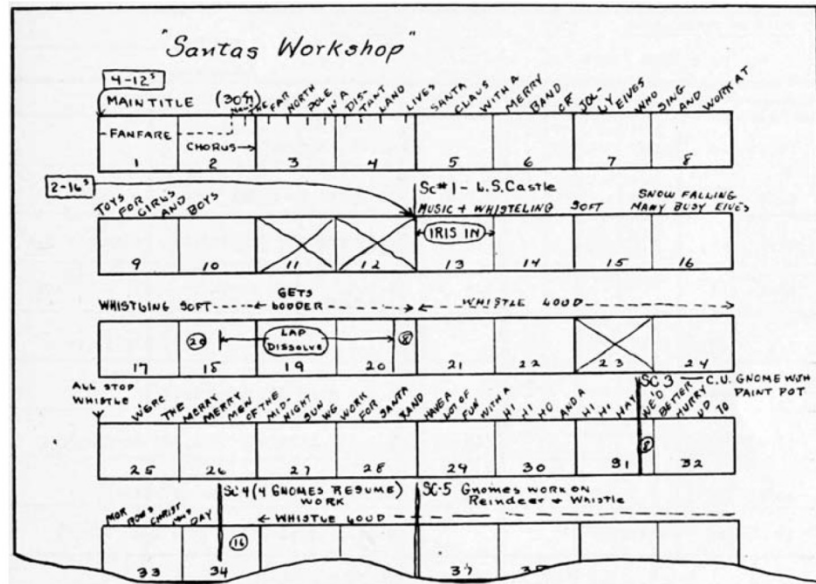


Figure 13 : Feuille à barres du film *Santa's Workshop*  
 Source : [JACOBS, 2015, p. 60]

On peut regarder comment dans la feuille sont incluses les paroles de la chanson et le nombre de photogrammes avec les indications de coupage<sup>146</sup>. Ce système, exclusif du studio Disney au début, a été rapidement assimilé par d'autres studios d'animation à l'époque, en devenant une forme de praxis énonciative étendue. Du côté du dispositif de réception, l'assimilation de l'effet de *mickey mousing* a été aussi immédiate, au point que cet effet a transposé du monde de l'animation vers celui du dispositif du cinéma en général, bien qu'on reconnaisse toutefois que la musique imitait le mouvement d'un personnage, ou qu'un bruit était remplacé par la musique d'un instrument. Daniel Goldmark raconte l'impact du film *Steamboat Willie* dans le public et les critiques à l'époque de sa première projection, le 18 novembre 1928 :

*Mickey réagit non seulement visuellement, mais aussi de façon audible à chaque nouvelle direction prise dans le récit. De même, on peut entendre aussi bien que voir ses exploits musicaux. Des comptes rendus contemporains indiquent que le film fait appel aux critiques*

<sup>146</sup> Pour une description du *mickey mousing*, voir (LUCENA JÚNIOR, 2002, p. 104-105).

*non pas parce que le film avait du son (...), mais plutôt à cause de son utilisation singulière de la musique et les effets synchronisés.* (GOLDMARK, 2007, p. 226)<sup>147</sup>

Goldmark retrace les antécédents de la synchronisation de la musique avec les « *cartoons* » avant l'apparition du cinéma sonore, pour montrer qu'il y avait déjà toute une préparation des publics pour recevoir avec enthousiasme le *mickey mousing*. L'auteur souligne le fait qu'à cet égard, l'animation a assimilé et développé avant tout autre cinéma les possibilités expressives du son synchronisé.

Mais tandis que le *mickey mousing* s'est répandu en Amérique du Nord (et très vite autour du globe), de l'autre côté de l'Atlantique d'autres animateurs et techniciens du son exploraient aussi les possibilités des différents techniques de l'enregistrement sonore, en donnant lieu à une autre tradition –et un autre dispositif, on va le voir– de synchronisation son/image. En raison de la nouveauté qu'il a introduite dans les modes de production audiovisuelle –et même dans la composition musicale–, nous voulons maintenant nous concentrer sur le deuxième exemple que nous venons de citer: le cinéma d'animation dit expérimental qui s'est concentré sur la recherche d'une synchronisation parfaite entre l'image et le son. On va donc maintenant examiner un ensemble de pratiques liées à une invention technique, un type de support et une nouvelle textualité audiovisuelle...

### **Le dispositif animatique du cinéma dit « expérimental »**

La distinction entre différents dispositifs animatiques est partiellement équivalente aux distinctions faites par les historiens et théoriciens de l'animation qui, dans nos termes, se concentrent selon le cas, sur différents niveaux de pertinence. Ainsi, par exemple, Maureen Furniss se déplace du niveau des textes filmiques animés (en parlant de formes et de styles, mais aussi en faisant des analyses concrètes) au niveau des pratiques de production (les techniques et les formats, selon les médias de distribution), et de réception (les audiences de l'animation), jusque dans le domaine des formes de vie : elle consacre un chapitre aux « régulations institutionnelles » où elle parle du « goût » : la régulation du contenu des émissions pour refléter un standard éthique établi, et du « contrôle », comme « l'accès –le

---

<sup>147</sup> “*Mickey reacts not only visually but also audibly to each new direction taken in the narrative. Likewise, we can hear as well as see his musical exploits. Contemporaneous reviews indicate that the film appealed to critics not because the film had sound (...), but rather because of its singular use of synchronized music and effects*”.

problème de qui est autorisé à voir un matériel donné » (FURNISS, 2007, p. 199). Ces standards d'émission reflètent –mais aussi produisent– des types de publics régulés institutionnellement.

Dans la même veine, le théoricien Paul Wells, dans son livre *Understanding animation* (WELLS, 2010) se déplace des réflexions sur « les styles et les approches » (chapitre 2), en passant par des descriptions des « stratégies narratives » (chapitre 3) et des figures rhétoriques (chapitre 4) jusqu'à une description des « audiences de l'animation » (chapitre 6), en prenant le cas, une fois de plus, de Disney et son influence dans les souvenirs d'enfance de l'audience de masse (WELLS, 2010, pp. 222 - 243).

Ces deux auteurs, avec des critères similaires, opposent aux formes industrielles de production, d'autres formes comme « l'animation expérimentale » (Wells, 2010, pp. 43 - 46)<sup>148</sup> ou « l'animation abstraite » (Furniss, 2007, pp. 249 - 262). Ces auteurs conviennent également de décrire alternativement le dispositif de production/création et le dispositif de réception, d'après l'expérience du spectateur, pour essayer de définir ces formes non « orthodoxes » ou « industrielles » d'animation. Ainsi, Furniss :

*Dans l'animation abstraite, il n'y a pas des personnages avec lesquels s'identifier, il n'y a pas de diégèse qui transporte le spectateur à un temps et un endroit différents et, quand l'animation est finie, le spectateur ne doit pas en avoir une « compréhension » complète de son sens comme il ou elle devrait l'avoir avec une structure narrative fermée. (FURNISS, 2007, p. 250)<sup>149</sup>*

Pour sa part, Wells affirme :

*Ce type d'animation est conçu pour résister à la configuration<sup>150</sup> selon laquelle souvent les audiences la regardent (c'est-à-dire, comme l'expression d'un personnage à travers le*

---

<sup>148</sup> L'abstraction étant un de ces traits principaux. Voir Wells, 2010, p. 43.

<sup>149</sup> “In abstract animation, there are no characters with which to identify, there is no diegesis to transport the viewer to a different time and place and, when the animation is over, the viewer does not have a complete ‘understanding’ of its meaning as he or she would with a closed narrative structure”.

<sup>150</sup> Par « configuration » Wells comprend le fait que l'animation dite orthodoxe présente des « figures », c'est-à-dire, « des personnes identifiables ou des animaux qui correspondent à ce que les audiences pourraient comprendre comme un être humain ou une créature orthodoxes, sans qu'il importe que son concept du design soit très coloré ou excentrique » (identifiable people or animals

*portrait d'un être humain ou d'une créature). L'animation expérimentale, par contre, redéfinit « le corps » ou résiste à l'utiliser comme une image illustrative. Les films abstraits sont plus préoccupés par le rythme et le mouvement en eux-mêmes et non par le rythme et le mouvement d'un personnage particulier. À cette fin, diverses formes et silhouettes sont souvent utilisées plutôt que des figures. (WELLS, 2010, p. 43)<sup>151</sup>*

Pour Wells l'animation abstraite se caractérise aussi par sa non-continuité spécifique, son intérêt pour l'évolution de la matérialité de son *médium*, la multiplicité des styles, la forte présence de l'artiste et sa puissante relation avec la musique (l'utilisation de sa dynamique). Ce sont ces deux derniers traits qui nous intéressent le plus, mais pour l'instant, surtout celui concernant la musique et le son :

*... on peut suggérer que si la musique pouvait être visualisée elle ressemblerait à des couleurs et à des formes en mouvement à travers le temps avec différents rythmes, mouvements et vitesses. Beaucoup de films expérimentaux cherchent à créer cet état, et comme je l'ai déjà suggéré, certains cinéastes perçoivent qu'il y a une relation psychologique et émotionnelle avec le son et la couleur qui peut être exprimée par la forme libre qui caractérise l'animation. Le son est important dans tout film d'animation, mais il y a une résonance particulière dans le film expérimental, comme ce dernier résiste souvent à avoir des dialogues ou des effets sonores clichés du cartoon, ou l'émotivité facile de certains types de musique. (WELLS, 2010, p. 46)<sup>152</sup>*

---

who corresponded to what audiences would understand as an orthodox human being or creature, despite whatever colourful or eccentric design concept was related to it) (Wells, 2010, p. 36).

<sup>151</sup> *“This kind of animation tends to resist configuration in the way audiences most often see it (i.e. as an expression of character through the depiction of a human being or creature). Experimental animation either redefines “the body” or resist using it as an illustrative image. Abstract films are more concerned with rhythm and movement in their own right as opposed to the rhythm and movement of a particular character. To this end, various shapes and forms are often used rather than figures”.*

<sup>152</sup> *“...it may be suggested that if music could be visualized it would look like colours and shapes moving through time with differing rhythms, movements and speeds. Many experimental films seek to create this state, and as I have already suggested, some film-makers perceive that there is a psychological and emotional relationship with sound and colour which may be expressed through the free form which characterizes animation. Sound is important in any animated film, but has particular resonance in the experimental film, as it is often resisting dialogue, the clichéd sound effects of the cartoon, or the easy emotiveness of certain kinds of music”.*

D'une façon similaire, le théoricien Patrick Barrès parle du cinéma d'animation qui fait partie des expériences artistiques et esthétiques des arts plastiques. Barrès met l'accent sur les pratiques de production/création des films et, surtout, sur la matérialité du médium :

*Il est nécessaire, parmi la grande variété des pratiques actuelles du cinéma d'animation, d'identifier celles qui se démarquent d'une visée exclusive de la représentation et qui rejoignent une démarche de création centrée sur les matériaux et les outils, les pratiques gestuelles et tout ce qui les dirige, notamment la saisie des « occasions opportunes », plus globalement vers une pratique de chantier. Elles résultent de l'engagement que prennent les artistes à renouveler les procédés du cinéma d'animation, en intégrant le décloisonnement des pratiques, à expérimenter les paramètres et les dispositifs impliqués dans la production d'images et d'images-mouvement, à explorer un univers plastique visuel et sonore (ouvert à une appréhension sensorielle tactile des images), finalement à repenser l'ensemble des conduites créatrices dans ce domaine. (BARRÈS, 2006, p. 19)<sup>153</sup>*

Comme on peut le constater dans ces citations, pour définir la particularité du cinéma d'animation dit « expérimental » ou « abstrait », il est presque inévitable de prendre en considération le travail créateur de l'artiste et ses pratiques. Et parler des pratiques créatrices implique de parler aussi des matériaux sur lesquels on travaille. L'intérêt, comme le fait Barrès, de rapprocher le cinéma d'animation des beaux-arts va dans ce sens. On peut dire qu'il s'agit d'un parcours du plan de l'expression descendant dans le sens de Fontanille : la pratique qui travaille sur un matériau pour produire un objet esthétique, est aussi la face matérielle d'un texte audiovisuel, et lui confère sa face formelle.

Néanmoins, dans le cas qu'on veut examiner ici, ce qui a permis l'émergence d'un nouvel ensemble de pratiques créatrices ainsi qu'une nouvelle forme de textualité audiovisuelle, a été une innovation technique, relative à l'objet-support. Il s'agit de la synthèse optique des sons.

### **La synthèse optique du son : un nouveau support pour de nouvelles pratiques**

Le concept du dispositif nous a placés dans un point de convergence entre la dimension matérielle du cinéma et sa dimension conceptuelle ou langagière. Dans le cas qui nous concerne maintenant l'accent est encore plus fortement placé sur ce point de convergence. L'invention technique dont on parle est, premièrement, l'inscription acoustique

---

<sup>153</sup> C'est moi qui souligne.

ou la *mise-en-écriture* du son, sa transcription, mais aussi la possibilité technique de manipuler et même de produire de nouveaux sons à partir de leur forme graphique ; et, finalement, la possibilité de « rephénoménaliser » le son inscrit ou « écrit » : le rendre audible, et en synchronisation avec l'image<sup>154</sup>. Il s'agit de tout un chapitre de l'histoire de la technologie audiovisuelle, où retentissent les noms des ingénieurs et inventeurs comme le finlandais Eric Tigerstedt<sup>155</sup>, les Allemands Josef Engl, Hans Vogt et Joseph Massolle<sup>156</sup>, ou

---

<sup>154</sup> Selon Thomas Y. Levin il est possible de tracer la généalogie de la transcription acoustique, en remontant à la découverte de Ernst Florens Friedrich Chladni (1787), le « père » de l'acoustique : « la découverte de Chladni du fait qu'une couche de poussière de quartz sur une feuille de verre pouvait, avec la vibration produite par un archet de violon, former des motifs distincts et réguliers ou des figures de ton (*Klangfiguren*), comme il les appelait, qui correspondent à des tons spécifiques, a effectivement démontré l'existence de traces visuelles des hauteurs, dont le caractère iconique-indiciel les différencie de façon sémiotiquement cruciale de tous les autres moyens conventionnels de notation de son » [*Chladni's discovery that a layer of quartz dust upon a sheet of glass would, when vibrated by a violin bow, form distinct and regular patterns or Klangfiguren (tone figures), as he called them, that correspond to specific tones, effectively demonstrated the existence of visual traces of pitches whose iconico-indexical character differentiated them in a semiotically crucial fashion from all other conventional means of notating sound.*] (LEVIN, 2006, p. 39).

Le moment suivant de cette généalogie a été la restitution du son comme trace visible, sujet de la phonographie. Des inventions comme le « phon-autographe » d'Edouard Léon Scott (1859), en passant par les enregistrements photographiques d'Edward L. Nichols et Ernst George Merrit, jusqu'à l'invention d'Edison en 1877, qui a rendu possible non seulement la visualisation du son, mais aussi sa re-phénoménalisation : « Ce qui était auparavant une graphématique visuellement accessible, mais silencieuse, de l'acoustique, était désormais capable tant de tracer que de rephénoménaliser le son, mais au moyen d'une inscription qui –en un geste de coquetterie media-historique– cache les secrets de sa spécificité sémiotique dans les recoins des rainures phonographiques » [*What had previously been a visually accessible but nonsounding graphematics of the acoustic was now capable of both tracing and rephenomenalizing sound, but by means of an inscription that—in a gesture of mediahistorical coquetry—hid the secrets of its semiotic specificity in the recesses of the phonographic grooves.*] (Levin, 2006, p. 40). Ce qui est intéressant pour Levin c'est l'utilisation de ces technologies dans la recherche linguistique.

<sup>155</sup> Eric Magnus Campbell Tigerstedt est le créateur de la première technologie du son sur pellicule. En 1914, il a présenté son film *Word and Picture* à Berlin, avec son invention, appelée à l'époque le « photomagnétophone ». Tigerstedt a obtenu le brevet allemand de la technologie du son sur pellicule en Juillet 1914, mais il l'a perdu après la Première Guerre mondiale.



les Américaines Theodore Case, Freeman Harrison Owens et Lee de Forest<sup>157</sup>. Ces inventeurs (et d'autres) ont développé une technologie de synchronisation son/image alternative au système du son enregistré sur un disque, qui offrait de nombreuses difficultés pour réaliser la synchronisation avec l'image<sup>158</sup>. Il s'agit d'une technologie que Philippe Langlois décrit comme suit :

*Pour enregistrer un son sur piste optique, qui est en fait une pellicule photosensible, il faut relier un microphone à l'appareil enregistreur (une tireuse optique) afin de recueillir les micro-impulsions électriques transmises par la membrane du microphone captant un signal acoustique de très faible intensité. Ces variations sont ensuite converties en images photographiques en actionnant la commande d'ouverture et de fermeture d'un faisceau lumineux qui expose la pellicule en fonction de la densité du signal transmis. Grâce à un jeu de prismes, la lumière converge en un point précis sur la pellicule – une petite ouverture verticale qui « impressionne » une portion de la piste optique photosensible – en défilant à la vitesse de vingt-quatre images par seconde soit environ quarante-trois centimètres. L'information sonore se trouve alors fixée sous la forme d'une image continue de l'onde sonore émise.*

*De la même manière, la lecture optique s'effectue sans aucun contact mécanique. Lorsque le film est projeté, la bande sonore défile devant le lecteur optique constitué d'un côté d'une source lumineuse et de l'autre d'une cellule photoélectrique. En défilant entre cet émetteur-récepteur de lumière, les ondes sonores inscrites sur la piste optique transforment les variations de lumière en impulsions électriques. La cellule est ensuite elle-même reliée à un amplificateur, de manière à ce que les ondes sonores visuelles soient rendues audibles.*

---

<sup>156</sup> Créateurs du système Tri-Ergon, breveté en 1919 et acheté par la Twentieth Century Fox en 1926. En combinaison avec les inventions de Theodore Case et Freeman Harrison Owens, cette société a développé le système Movietone, un des plus popularisés dans les premiers temps du cinéma parlant.

<sup>157</sup> Créateur du système du son optique appelé le « Phonofilm » en 1919, de Forest a travaillé avec Case et Owens, mais ils se sont séparés. Après cette rupture, il a vendu aussi ses inventions à la Fox. Selon Russell Lack, en 1924 plus de trente théâtres américains étaient déjà prêts à intégrer le système Phonofilm. Néanmoins, le système final adopté par la Fox a été le « Movietone », le résultat de l'union de la technologie du « Phonofilm » et le système Tri-Ergon des allemands. Voir (Lack, 1997).

<sup>158</sup> Bien que ces systèmes ont réussi à produire une synchronisation presque parfaite. Le plus chanceux de ces systèmes a été le « Vitaphone », un graveur de disques couplé à la caméra et lecteur de disques couplé au projecteur. En effet, le Vitaphone est le système utilisé dans le film *Le Chanteur de jazz* (Alan Crossland, 1927), célèbre comme le premier « Talkie ».



*L'enregistrement du son cinématographique relève donc, tout comme l'image, d'un subtil contrôle de la lumière.*

*La caractéristique technique spécifique du son cinématographique est donc d'exister sous la forme d'une image photographique reproduisant le graphisme d'une onde sonore sur la piste optique de la pellicule. Le lien technique entre son et image se resserre ici jusqu'à figer le son dans une représentation visuelle telle une modulation en négatif et le noir et blanc de la forme d'onde sonore enregistrée et ce, quelle que soit la source enregistrée. Le support du son au cinéma se limite donc à une représentation purement visuelle du son sous une forme photographique. A l'intérieur même du projecteur, le son y est projeté tout comme l'image, c'est-à-dire à partir des variations de lumière produites par la densité photographique de la forme d'onde. En rejoignant cette dimension optique, le son s'intègre désormais totalement au dispositif cinématographique sous la forme d'une information visuelle contenant tous ses paramètres : la hauteur, la durée, l'intensité, le timbre. (LANGLOIS, 2012, pp. 145 - 146)*

Le système du son sur pellicule a permis une technologie d'enregistrement et de projection où le son et l'image sont inscrits dans le même support matériel, ce qui permet aussi le travail parallèle sur les deux matières de l'expression filmique. Dans ce sens, l'invention de la synthèse optique du son n'est que la première partie de ce parcours, l'histoire de la synchronisation développée après surtout dans le cinéma PVR et les films narratifs dits institutionnels. La seconde partie a à voir avec *la production graphique du son* à l'intérieur du cinéma d'animation, où « les cinéastes adapteront ainsi ce genre de techniques visuelles propres à son fonctionnement, notamment en s'émancipant de la caméra. Le film d'animation devient ainsi essentiellement un art du support visuel et sonore » (LANGLOIS, 2012, p. 147).

Ce type d'animation a reçu le nom d'animation « directe »<sup>159</sup> et, d'après plusieurs auteurs, c'est la technique où le réalisateur audiovisuel a été plus proche de l'artiste dans la mesure où il/elle a un contrôle total sur sa création, ce qui renforce l'argument de Barrès. Cette étape introduit aussi une nouvelle manière de créer qu'on peut décrire d'après le niveau

---

<sup>159</sup> Pour une définition et des exemples de l'animation directe, voir (Furniss, 2007, pp. 40 - 45). Les expériences des russes et des allemands dont on va parler dans les prochaines pages sont, néanmoins, différentes, parce qu'il y a besoin d'une caméra pour photographier le son dessiné sur cartons ou papier, de manière qu'il soit enregistré sur la pellicule. Il s'agit donc de son « dessiné », mais pas d'animation « directe » du son. Dans le paragraphe consacré à Norman McLaren, on va parler aussi de ce type production sonore et de ces différences avec le son optique dessiné sur papier. Il s'agit de deux pratiques sémiotiques proches, mais différentes.

des « formes de vie » : parce que ce type de travail créatif implique l'appropriation esthétique du dispositif de production par une volonté artistique, mais avec les connaissances technologiques des ingénieurs. Il s'agit de l'émergence d'un nouveau travailleur-créateur ou artiste-technicien propre au cinéma, mais qui dans le domaine de l'animation prend une forme particulière<sup>160</sup>.

On a donc un nouveau type de dispositif animatique de production, avec un support (la pellicule), une pratique (l'animation directe du son et de l'image) et de nouvelles formes de textualité : les films *synchromiques*, pour reprendre la formule de Langlois. Mais on peut dire aussi qu'on est aussi en face de nouveaux types de *codification du son* ou, en d'autres termes, des unités sonores minimales, des *signes sonores*. Pour traiter cette question, il est nécessaire à regarder le travail des pionniers du traitement graphique du son. Dans cette histoire il y a au moins trois lignes généalogiques qui se disputent cette invention : les Russes Arseny

---

<sup>160</sup> Un exemple d'intérêt est la Société Leonardo da Vinci pour les Arts et les Sciences, fondée par des artistes du mouvement futuriste russe, y compris plusieurs des animateurs qui ont développé le son synthétique dans ce pays. Un d'eux, Evgeny Sholpo, déclara : « en assumant le nom de da Vinci, nous avions à l'esprit strictement l'idéologie scientifique et esthétique (et plus scientifique qu'esthétique) de cet homme brillant. Avec foi en la puissance de la science et des mathématiques, la direction vers la connaissance objective des lois "mystérieuses" de l'art –ont conduit la Société à prendre ce nom historique... Nos regards ont été mis sur la révolution dans la théorie de la musique et de la technologie, issu d'une fusion de l'art et de la science... » [*...in assuming da Vinci's name we had in mind strictly the scientific and aesthetic (and more scientific than aesthetic) ideology of that brilliant man. Faith in the power of science and mathematics, the drive toward the objective cognition of the "mysterious" laws of art -these compelled the Society to take on that historic name... Our sights were set on the revolution in music theory and technology, stemming from a fusion of art and science...*]. Cité dans (Izvolov, 2014, pp. 63 - 64). En effet, on peut identifier deux attitudes différentes en face de la création artistique, qu'on va retrouver aussi ci-dessous chez les réalisateurs allemands du « son dessiné » : une voie dite « romantique » associée à la liberté créatrice, le dessin direct sur la pellicule et la plasticité du son (celle d'Oskar Fischinger), et une voie dite « scientifique » ou « technique » comme celle des russes, basée sur un système contrôlé d'unités qui peut être répliqué (celle de Rudolph Pfenninger).

Avraamov, Evgeny Sholpo, Boris Yankovsky et Nikolai Voinov ; les Allemands Oskar Fischinger et Rudolph Pfenninger ; et l'anglais Eric Allan Humphriss<sup>161</sup>.

Il y a aussi une autre origine –au moins comme idée– dans un article de Guido Bagier, le directeur artistique du département du Son de l'UFA (Universum Film AG, principale société de production cinématographique allemande de la république de Weimar), publié en 1926 dans le journal de la musique, *Musikblätter des Anbruch*. Dans cet article, intitulé « Musik und Maschine » (la musique et la machine), Bagier présentait le système du son optique aux musiciens et compositeurs, en suggérant la possibilité de manipuler les traces du son ou même de les fabriquer à la main pour construire une nouvelle forme de notation musicale et, par conséquent, une nouvelle méthode de composition. Voir (Patteson, 2015, pp. 82 - 83). Néanmoins, dans ce cas il s'agit de profiter les nouvelles possibilités de la technologie du son sur la pellicule pour la composition musicale. Cette voie a été continuée par Robert Beyer, un ingénieur du son et musicien qui est devenu un activiste de la production musicale à travers la technologie du son optique. D'après Patteson, pour Beyer,

---

<sup>161</sup> Humphriss était un physicien et ingénieur britannique qui travaillait pour la British International Film Corporation, qui a fait la première production de « parole synthétique » en manipulant les sons directement sur ces courbes graphiques sur la pellicule en 1931. Les nouvelles ont parus dans le New York Times, avec la description suivante: « Discours synthétique démontrée à Londres: un ingénieur a créé une voix qui n'a jamais existée » (Synthetic Speech Demonstrated in London: Engineer Creates Voice which Never Existed). Voir (LEVIN, 2006, p. 33), et aussi (PELLETIER, 2009). Dans les histoires sur les origines du son synthétique « dessinée ou écrit » le nom de l'artiste László Moholy-Nagy est toujours mentionné. En effet, cet artiste a été intéressé sur les différentes techniques de production sonore à partir des formes visuelles. Néanmoins, Moholy-Nagy a reconnu et souligné le rôle pionnier des cinéastes et inventeurs allemands, Fischinger et Pfenninger, qui ont été la source de ses articles et expérimentations. La contribution de Moholy-Nagy au cinéma du son graphique est le film *Tönendes ABC* (*Son ABC*, 1933), où il utilise tout type de signes, symboles, lettres et même ces empreintes digitales dans la bande-son, pour produire toute sorte de bruits. Voir (Levin, 2006, pp. 44 - 50).

Dans un article intitulé *Son Animé sur Pellicule : les débuts*, Norman McLaren mentionne également à Ernst Toch, théoricien allemand qui « suggéra l'écriture directe du son sans le musicien ». Il parle aussi de Jack Ellit, réalisateur australien qui, en Angleterre, « suivit une ligne de recherche semblable à celle de Pfenninger et en 1933 » et qui « ouvrit les routes au système du dessin du son directement sur la pellicule, sans utiliser la caméra » (McLaren, *Son Animé sur Pellicule. Les Débuts*, 1997, p. 155).

*Le cœur de la nouvelle technique de composition réside dans le processus de «chimie sonore» à travers laquelle le matériel enregistré est transformé en figures, formes et tropes acoustiques, des unités syntaxiques du nouveau langage musical, comparable par sa fonction avec les tons et les thèmes de la technique de composition traditionnelle. Le compositeur «atomise» et «dynamise» le matériel enregistré, en le décomposant et le reconstruisant dans de nouvelles configurations perceptives. Ces nouvelles figures sont ensuite assemblées en unités de composition plus grandes via la technique de montage. Grâce à ce processus, la génération de la tonalité électrique prend un sens productif par opposition au sens reproductif. (PATTESON, 2015, p. 102)<sup>162</sup>*

Au-delà du domaine de l'animation, qui est notre principal intérêt ici, le son optique a ouvert aussi un champ de recherche intéressant pour la sémiotique du son en général et de la musique en particulier. Pourtant, ces recherches étaient partagées par les réalisateurs qui cherchent la synchronisation parfaite entre l'image animée et le son dessiné<sup>163</sup>.

Dans le cas des Russes, Nikolai Izvolov cite le témoignage de Mikhail Tsekhanovsky, cinéaste et animateur, sur une rencontre où aurait émergé l'idée du « son dessiné » :

*Nous étions assis dans le studio –moi, E.A. Sholpo, que j'avais invité à être mon assistant, et l'artiste-animateur M.M. Tsekhanovsky (...). Avec un immense intérêt, nous utilisions une loupe pour examiner la première et fraîche impression d'une bande sonore, encore humide, qui venait d'arriver du laboratoire. « C'est intéressant. Si vous dessiniez une figure égyptienne ou un dessin ancien grec dans la bande-son, pourrions-nous entendre de la musique archaïque jusqu'alors inconnue? » [a dit Tsekhanovsky]. Sholpo et moi, nous avons apporté*

---

<sup>162</sup> “The heart of the new compositional technique lay in the process of “sonic chemistry” through which the recorded material is transformed into acoustic figures, forms, and tropes—the syntactic units of the new musical language, comparable in function to the tones and themes of traditional compositional technique. The sound-film composer “atomizes” and “dynamizes” the recorded material, breaking it down to its component parts and reconstructing it into novel perceptual configurations. These new figures are then assembled into larger compositional units via montage technique. Through this process, electric tone generation takes on a productive as opposed to reproductive meaning”.

<sup>163</sup> On peut placer ici l'origine du montage sonore, en le comprenant comme une troisième pratique sémiotique rendue possible par le son optique. Dans ce cas, les sons sont enregistrés avec des microphones, mais, une fois transformés en pellicule, ceux-ci peuvent être coupés, collés et mis en séquences de nouvelles façons.

*sa vision de retour sur terre. Comme l'ornement lui-même a une forme fortement périodique, en fonction de sa forme nous entendrions seulement des tons simples d'un timbre ou d'un autre. Qu'ils seraient « grecs » ou « égyptiens » est difficile à dire, mais ils ne seraient certainement rien qui ressemble à une mélodie...*

*Mais le mot avait été prononcé. L'idée de reproduire une bande sonore synthétique « artificielle » sur la pellicule avec toutes ses possibilités, cette brillante idée avait pris fermement notre main à tous. (IZVOLOV, 2014, pp. 45 - 46)<sup>164</sup>*

Effectivement, chacun de ces créateurs a pris son propre chemin pour produire des sons synthétiques. Le premier à produire des sons dessinés a été Arseny Avraamov en 1930. Les autres trois ont même créé des dispositifs techniques pour faciliter le dessin du son sur la pellicule. Sholpo a créé le « variophone » (appelé par ces collègues le « sholpograph »). Boris Yankovsky a construit le « vibroexponator » avec la particularité de travailler sur des différences de timbre. Nikolai Voinov a créé une espèce de peigne de papier qui fonctionnait comme modèle pour dessiner les photogrammes acoustiques. Avec cette technique, appelée « son de papier », Voinov a produit le film *Vor (Voleur, 1934)*<sup>165</sup>.

---

<sup>164</sup> “The three of us were sitting in the studio—myself, E. A. Sholpo, whom I had invited to be my assistant, and artist-animator M. M. Tsekhanovsky (the maker of the first Soviet sound animation film, “The Postal Service” [Pochta, Tsekhanovskii and Timofeev, 1929], based on the work of Samuil Marshak). With immense interest we were using a magnifying glass to examine the very first, fresh print of the soundtrack, still moist, which had just arrived from the lab. Tsekhanovsky enthused about the beauty of the ornamental waveform traced on the film. He fantasized:

“Interesting, if you were to trace an Egyptian or ancient Greek design on the soundtrack—would we hear some hitherto unknown archaic music?” Sholpo and I brought his fantasy back down to earth. As the ornament itself is strongly periodical in form, depending on its shape, we would hear only single tones of one timbre or another. Whether they would be “Greek” or “Egyptian” is hard to say, but there would certainly be nothing resembling a melody... But the word had been spoken. The idea of reproducing a synthetic, “artificial” soundtrack on the film strip with all its brilliant possibilities—this idea had taken firm hold of us all”.

<sup>165</sup> On trouve les descriptions de ces inventions avec plus de détails dans (Izvolov, 2014). En regardant le film *Vor*, on trouve de nombreuses similitudes avec les films de McLaren, particulièrement avec *Synchromie*. Izvolov nous rappelle que le cinéaste écossais connaissait les expériences des animateurs soviétiques. On peut regarder un petit échantillon de ces expériences sur le lien suivant :

<https://www.youtube.com/watch?v=Z7Zb4rso82M>

Dans le même temps, et sans savoir ce que les Russes faisaient, les Allemands Oskar Fischinger et Rudolph Pfenninger travaillaient aussi sur le son dessiné. Néanmoins leurs buts et techniques étaient très différents. Fischinger venait de la tradition du premier cinéma abstrait, où il cherchait « la musicalité des formes graphiques en mouvement » (Levin, 2006, p. 51). À partir de 1932, il a entrepris ses expériences avec le son dessiné avec l'idée de trouver une correspondance ou isomorphisme entre les formes visuelles et les sons qui leur sont associés. Ses premiers films (les *Expérimentations avec le Son synthétique*) « ont été composés par « motifs dessinés en papier avec plume et encre et photographiés directement sur la marge du film réservé à la bande-son » (Levin, 2006, p. 51). Fischinger a défendu l'idée que ce type de production sonore avait rendu au cinéaste le contrôle total sur sa création, en restaurant sa puissance artistique.

Trois ans plus tôt (en 1929 environ) l'ingénieur et animateur Rudolf Pfenninger a entrepris la même sorte d'expériences avec le son dessiné sur la pellicule. Néanmoins sa quête était plus pragmatique et sa connaissance de la technologie plus étendue, donc ses résultats ont été plus effectifs. D'après Pfenninger lui-même :

*...dans mes travaux j'ai étudié la nature des sons à l'aide d'appareils appelés oscillographes. J'ai en fait décomposé les images sonores en sons basiques. Il ne me manquait plus qu'à pouvoir contrôler ces analyses. Il m'est ainsi venu l'idée de dessiner chacune des oscillations sur des bandes de papier, de photographier le tout et d'écouter le résultat sur un projecteur. Maintenant que j'avais la possibilité d'écouter des oscillations sonores dessinées par mes soins, j'ai essayé de dessiner de nouvelles formes d'après mes propres calculs afin de produire des sons qui n'existent pas dans la nature (...) les sons obtenus avaient un caractère fantastique, absolument nouveau<sup>166</sup>.*

Dès le début, le but de Pfenninger a été la production des sons qui ne dépendent pas des sources du monde mais qui étaient produits directement sur la pellicule<sup>167</sup> : les « tons provenant de nulle part ». Les différences entre ces deux projets peuvent également être vues

---

<sup>166</sup> Déclaration de Pfenninger dans une interview avec Helmuth Renar. Cité dans (LANGLOIS, 2012, p. 153).

<sup>167</sup> En fait, selon Levin la légende veut que c'est le manque d'argent pour financer l'accompagnement sonore et musical de ses films qui a conduit Pfenninger à concevoir le son de cette façon.

dans ces titres : les *Tönende Ornamente (Ornements sonores)* de Fischinger vs les *Tönende Handschrift (Écriture sonore)*. En effet, la différence la plus grande n'est pas seulement technique, ni esthétique, mais sémiotique :

*Pfenninger a mis l'accent sur le développement technologique d'une nouvelle forme d'écriture acoustique, une sémio-pragmatique du son dont la fonction était de libérer la composition des contraintes des instruments musicaux existants et des conventions de notation régnantes. Contrairement à Fischinger, qui a commencé avec des formes graphiques et ensuite a exploré le genre de sons qu'ils produisent, l'objectif principal de Pfenninger portait sur l'acoustique, dans une tentative d'établir quelle était la forme d'onde précise pour reproduire un son spécifique à volonté. Malgré l'attrait visuel potentiel de leurs formes d'ondes sinusoïdales, les courbes de Pfenninger ne sont décidément pas des ornements, mais sont plutôt, comme de nombreux critiques l'ont noté à juste titre, des « modèles ou caractères de presse » ; notamment, des entités sémiotiques qui peuvent être combinées pour produire des sons d'une façon 'linguistique' –c'est-à-dire, d'une façon complètement technique et régie par des règles–. Contrairement aux courbes de Fischinger, qui étaient continues, les courbes de Pfenninger étaient des unités discrètes. En effet, ce qui permet de différencier les deux projets de manière la plus succincte est le fait que, tandis que Pfenninger aurait pu (au moins en théorie) utiliser sa méthode pour reproduire chaque son émis par les ornements de Fischinger, le contraire n'est pas évidemment le cas. Il n'est donc pas surprenant que les premiers critiques aient insisté à juste titre sur le fait que l'invention de Pfenninger n'était pas une pratique ornementale autant qu'elle était une nouvelle technique de notation acoustique, allant même jusqu'à prétendre qu'il était en train de « construire un appareil ressemblant une machine à écrire qui, au lieu de lettres, mettra ensemble des signes-vagues en succession ».* (LEVIN, 2006, p. 58)<sup>168</sup>

---

<sup>168</sup> “Pfenninger focused on the technological development of a new form of acoustic writing, a semio-pragmatics of sound whose function was to liberate composition from the constraints of both the extant musical instrumentarium and reigning notational conventions. Unlike Fischinger, who began with graphic forms and then explored what sort of sounds they produced, Pfenninger’s primary focus was on the acoustic, in an attempt to establish what the precise wave form is that would allow one to re-produce a specific sound at will. Despite the potential visual appeal of their sine-wave forms, Pfenninger’s curves are decidedly not ornaments but are rather, as numerous critics have rightly noted, “templates or print-types”<sup>63</sup>; that is, semiotic entities that can be combined to produce sounds in a linguistic—which is to say, thoroughly technical and rule-overned—manner. Unlike Fischinger’s curves, which were continuous, Pfenninger’s were discrete units. Indeed, in what is perhaps the most



Dans ce sens, le projet de Pfenninger est plus proche des expériences des Russes, qui ont réussi à créer des machines capables d'écrire ce type de sons synthétiques. En fait, l'invention de Pfenninger a révolutionné la conception traditionnelle de la musique et sa relation avec le son en général. D'une part, elle a changé le statut du son enregistré, en marquant une autre différence sémiotique avec Fischinger : « Si la production synthétique du son de Pfenninger a détruit effectivement la logique de l'indexicalité acoustique qui a été la base de tout son enregistré antérieur, elle a également exposé le résidu iconique-indexical des expériences de Fischinger » (LEVIN, 2006, p. 58). D'après Levin, la conception sémiotique des courbes sonores de Pfenninger était basée sur les propriétés « arbitraires » de la cellule de sélénium qui était à la base de son système, tandis que les ornements de Fischinger ont été conçus comme des signes « motivés », sur la base de la similitude.

Pour récapituler, on peut situer les différentes étapes de la découverte du « son dessiné » dans le cadre théorique présenté au début. Ainsi, l'invention du son optique appartient au *dispositif matériel de production* et elle est le résultat des pratiques des techniciens et des inventeurs, recueilli dans les stratégies des entreprises de production cinématographique pour trouver la forme de synchronisation la plus précise (et plus la attrayante pour le public) entre l'image et le son au cinéma en général. Cette recherche a conduit à l'émergence d'un nouvel objet-support : la bande sonore sur la pellicule. La découverte de la possibilité de produire de nouveaux sons en dessinant des formes qui pourraient être placées sur la pellicule directement ou à travers la photographie est une pratique créative particulière avec deux faces : une face matérielle de manipulation de l'objet-support, et une face langagière de création (dans les premières expériences) ou d'utilisation d'un code de notation, pour produire des sons organisés et synchronisés avec l'image. Les différences entre les procédures de Pfenninger et de Fischinger apparaissent à ce niveau : il s'agit de deux formes sémiotiques différentes qui utilisent le même support, la technique de Pfenninger étant la plus formalisée.

---

*succinct manner of differentiating the two projects, while Pfenninger could (at least in theory) have used his method to re-produce every sound made by Fischinger's ornaments, the opposite is obviously not the case. Thus it is no surprise that from the start critics rightly insisted that Pfenninger's invention was not an ornamental practice as much as it was a new technique of acoustic notation, even going so far as to claim that he was in the process of "constructing a contrivance resembling a typewriter which, instead of letters, will set together sign waves in succession".*



En outre, ce même support a été utilisé pour la production purement musicale : le son synthétique est devenu un nouvel instrument, au début de la musique électronique : le même support, mais une pratique différente. Dans ce sens, c'est la pratique qui organise les deux dimensions du dispositif, la dimension matérielle et la dimension « mentale », dans la terminologie de Sirois-Trahan.

Les résultats concrets de cette première étape sont partiels et fragmentaires. De nombreux films ont été perdus<sup>169</sup> et la plupart des autres sont des exercices ou des exemples de la technique plus qu des œuvres terminées. Néanmoins d'autres artistes, animateurs et cinéastes ont continué à explorer ces techniques, en produisant de véritables œuvres artistiques. Aux États Unis, les frères John et James Whitney ont développé une œuvre solide, en arrivant à créer leur propre système de production du son synthétique, à partir des pendules, un système « assez complexe qui permet de graver une musique directement sur la piste optique, dans une relation audiovisuelle et une synchronisation directe » (LANGLOIS, 2012, p. 167).

Après eux, certains artistes ont exploré les techniques du « dessin sonore ». Cionno Carpi, Barry Spinello et Alexandre Vitkine dans les années soixante, Liz Rodhes, Steve Farrer ou Pierre Rovère, parmi d'autres dans les soixante-dix...<sup>170</sup>. Tous ces réalisateurs ont écrit (presque littéralement) un chapitre parfois oublié de l'histoire du cinéma, mais aussi de la musique. Dans les termes de Langlois :

*La synthèse optique a permis de générer des sons qui n'appartiennent à aucune réalité instrumentale si ce n'est celle de l'abstraction sonore pure. Aussi bien chez les Russes que les Allemands, l'émergence de l'animation sonore correspond à un but commun non concerté,*

---

<sup>169</sup> À cet égard, un épisode malheureux nous permet de regarder le statut sémiotique de l'objet-support. Les films du réalisateur russe Arseny Avraamov ont été brûlés après sa mort, et perdus par l'histoire : « Son fils joue à la guerre avec ses amis en transformant la pellicule argentique en fumigène, en la découpant, en la chiffonnant ensuite dans des petits morceaux de journal avant d'y mettre le feu. Cela devait très certainement produire de très jolies fumées colorées servant pour ses jeux de guerre avec ses amis. Innocemment, c'est l'œuvre de son père qu'il détruisait » (Langlois, 2012, p. 159). La pellicule, comme le papier, est un simple objet pour celui qui ne connaît pas –ou ne perçoit pas- le « texte audiovisuel » inscrit sur sa matérialité.

<sup>170</sup> Voir (Langlois, 2012, pp. 394 - 400), pour une liste exhaustive et détaillée des artistes qui ont travaillé avec le son optique dans la seconde moitié du XXe siècle.

*enjeu de toutes les fascinations à travers la conquête d'une précision de la machine capable de surpasser les potentialités humaines notamment dans la précision rythmique. Grâce à ce mode de production sonore marginal, certains cinéastes d'animation sont parvenus à créer une complémentarité audiovisuelle inégalée. Ce corpus d'œuvres tout à fait à part dans la sphère du cinéma d'animation constitue probablement l'une des formes les plus abstraites, les plus abouties, les plus absolues de ce genre cinématographique. La synthèse optique représente sans aucun doute la méthode, de manière précoce, la plus novatrice sur le plan technique en matière de création sonore. Grâce aux cinéastes synesthésistes –Fischinger, les Whitney, McLaren- l'utilisation des sons synthétiques évolue vers une véritable création musicale. (LANGLOIS, 2012, p. 176)*

Et le nom le plus connu parmi ceux qui sont mentionnés par Langlois est, sans aucun doute, celui de Norman McLaren. Le réalisateur écossais/canadien a développé l'un des projets les plus vastes et systématiques dans le champ des relations image-son. Il est aussi l'auteur de ce qui est pour un grand nombre de critiques et historiens du cinéma, le film le plus complet et complexe réalisé avec la technique du son synthétique : *Synchronie* (1971). Maintenant, nous voulons aborder ce projet, qui nous permettra de regarder les dimensions et possibilités de ce dispositif animatique particulier.

### **Norman McLaren et la pratique sémiotique du dessin sonore**

*Avec le son dessiné, des formes, des teintes et des espaces deviennent rythmes, timbres, hauteurs de son et les barres comptées portées sur la piste sonore du film, mètres de sons modulés (...) En écrivant le son directement sur la pellicule ou en photographiant des éléments étalonnés, en modifiant l'équilibre des timbres, la dimension des silences McLaren, prend la musique à son niveau le plus général et le plus moderne : celui d'une ARCHITECTURE GÉNÉRALE DES FRÉQUENCES.*

*André Martin<sup>171</sup>*

En termes des possibilités expressives ouvertes par la technologie du son optique, la capacité de « montrer le dessin sonore » en simultanéité avec son expression acoustique est une des plus explorées par les animateurs. Les premiers films de Voinov montraient les

---

<sup>171</sup> (Martin, 2000, p. 244).

bandes de papier avec les dessins et découpages des ondes en même temps qu'on écoutait les traductions sonores de ces bandes. Cette possibilité, ouverte par une nouvelle technologie, a permis aussi l'émergence d'un autre parcours esthétique : la recherche, non seulement de la synchronisation parfaite, mais de l'*équivalence* parfaite entre la bande image et la bande-son. Comme le dit Langlois :

*La synthèse optique offre donc la possibilité de restituer fidèlement l'image du son et dans certains cas, quelques cinéastes parviennent à accomplir une relative synesthésie objective. Dès lors, la représentation auditive de l'image et la représentation visuelle du son sont possibles. Les synesthésies connaissent un nouvel élan à travers la représentation exacte et réciproque de l'image d'un son reporté sur la piste optique.* (LANGLOIS, 2012, p. 144)

La théoricienne et historienne du cinéma d'animation, Maureen Furniss donne une place spéciale aux pratiques productives dans ce domaine. Elle identifie le travail de Norman McLaren comme faisant partie de toute une tradition cinématographique :

*Norman McLaren était un autre des nombreux individus qui ont expérimenté avec la relation entre la musique et l'image animée. Un de ses intérêts principaux était la façon dont les différents composants sensoriels d'un film peuvent être considérés comme interdépendants ; lui, comme Sergei Eisenstein, Oskar Fischinger et bien d'autres artistes, a été intrigué par le concept de synesthésie, ou le chevauchement des sens.* (FURNISS, 2007, p. 88)<sup>172</sup>

Dans le chapitre treize de son ouvrage *Animation Aesthetics*, sur le rôle du son et de la musique dans la conception structurelle de plusieurs films dits « abstraits », Furniss parle à nouveau de McLaren, et elle cite une déclaration très intéressante de l'animateur :

*Si l'on est à la fois artiste et musicien, les chances qu'on devienne animateur sont beaucoup plus élevées, parce qu'on est intéressé par le mouvement – l'ensemble des flux et flots de ce qui se passe. La musique est organisée en termes de petites phrases, grandes phrases, des*

---

<sup>172</sup> “Norman McLaren was another of the many individuals to experiment with the relationship between music and animated imagery. One of his central interests was the way that different sensory components of a film could be seen as interdependent; he, like Sergei Eisenstein, Oskar Fischinger and many other artists, was intrigued by the concept of synesthesia, or the overlapping of the senses”.

*expressions, des mouvements entiers et ainsi de suite. À mon sens, l'animation est le même genre de chose*<sup>173</sup>.

Tel est, peut-être, le sens du « dispositif animatique » proposé par Philip Brophy : peu importe s'il s'agit du cartoon hollywoodien de Disney ou Warner, du cinéma plastique ou abstrait des avant-gardes européennes ou même des animations numériques d'aujourd'hui ; il y a toujours dans l'animation une volonté (ou au moins la possibilité) d'exprimer une seule structure rythmique pour les deux bandes (visuelle et sonore), bien que le degré de correspondance soit variable d'un film à l'autre. Et, particulièrement dans le dispositif de l'animation directe avec du son dessiné, cette volonté atteint ses plus hauts degrés de précision et versatilité.

Dans le cas de McLaren, cette volonté s'exprime de façons différentes. Terence Dobson a identifié cet intérêt dans un stade très précoce de sa jeunesse, quand il a lu le livre d'Adrian Bernard Klein, *Colour-music : the Art of Light*, de 1926, dans lequel l'auteur revendique le rôle des peintres du « synchronisme », un mouvement pictural du début du XX<sup>e</sup> siècle basé sur l'idée que la couleur et le son sont des phénomènes similaires et que la composition chromatique d'un tableau peut être organisée de la même manière que les compositeurs organisent les notes d'une symphonie<sup>174</sup>.

Tout au long de son travail, McLaren a exploré presque toutes les techniques et procédures d'assemblage image/son à son époque. Des films basés sur pièces musicales préexistantes, en passant par des films avec de la musique originale orchestrée ou jouée pour être enregistrée (composée ou pas pour le réalisateur) jusqu'à ses expériences avec le son optique. Dans tous les cas, il s'agissait de chercher à produire une structure commune :

*Norman McLaren structurait souvent ses films avec de la musique à l'esprit, même si la bande-son a été choisie ou créée après que l'animation soit complète; il a créé des*

---

<sup>173</sup> “If a person's a static artist and a musician, the chances that he or she will be an animator are much higher, because he's interested in motion –the whole flux and flow of what's happening. Music is organized in terms of small phrases, bigger phrases, sentences, whole movements and so on. To my mind, animation is the same kind of thing”. Norman McLaren, cité dans (FURNISS, 2007, p. 255).

<sup>174</sup> Il y a aussi une troisième édition plus longue et avec un autre titre, *Coloured Light : an Art Medium*, publiée en 1937. Sur l'influence de Klein et ces œuvres sur McLaren, Voir (DOBSON, 2006, pp. 27 - 32).

*mouvements qui ont eu lieu en battements rythmiques, de sorte qu'il était relativement facile de composer une partition pour compléter les visuels et post-synchroniser une bande-son. Donald McWilliams dit que la «planification des mouvements en termes de rythmes musicaux» était une «activité normale pour Norman » (FURNISS, 2007, p. 256)<sup>175</sup>.*

En ce qui concerne le son optique ou le son « dessiné », cette technique est habituellement associée à McLaren. Et bien qu'il soit loin d'être son créateur, c'est vrai qu'il l'a utilisée et développée plus que tout autre. Dans un programme enregistré en Allemagne en 1951, McLaren explique ces techniques :

*À l'Office National du Film du Canada à Ottawa, je suis parvenu récemment à produire du son synthétique au moyen de deux procédés. Le premier d'entre eux consiste à dessiner directement sur la pellicule vierge de 35 mm avec un crayon et un peu d'encre. Aucun autre appareil n'est nécessaire. On peut écouter le son en faisant se dérouler le film devant un lecteur sonore. Dans la plupart de cas, le dessin original est utilisé comme négatif pour tirer les copies (...) (MCLAREN, 1997, p. 158).*

Toutes les expériences de « son dessiné » directement ne sont rien d'autre que de brefs essais aux rythmes percutants. Mais il existe un deuxième procédé qui permet de créer une musique synthétique continue. On dessine la forme d'une onde sonore sur un bout de carton et on la photographie avec une caméra normale pour l'animation dans une position choisie pour qu'elle apparaisse sur la bande sonore au bord de la pellicule. C'est ce procédé qu'utilise l'Allemand Pfenninger dans ses expériences avec le son et c'est celui qu'on utilise toujours plus au Canada, comme support musical des dessins animés (MCLAREN, 1997, p. 159).

Selon Terence Dobson, bien que la technique des cartons dessinés et photographiés soit attribuable à Pfenninger, Avraamov et Voinov, comme on a déjà vu, l'origine du son dessiné ou rayé directement sur la pellicule est plus discutable : peut-être assignable à Moholy-Nagy (pour son film *Son ABC* de 1933) ou au musicien australien Jack Ellitt, qui travaillait initialement avec la technique de Pfenninger, et qui changea et commença à

---

<sup>175</sup> *“Norman McLaren often structured his films with music in mind, even when the soundtrack was chosen or created after the animation was complete; he created movements that occurred in rhythmic beats, so it was relatively easy to compose a score to complement the visuals and post-synchronise a soundtrack. Donald McWilliams says that the 'planning of movements in terms of musical rhythms' was a 'normal activity for Norman'”.*

dessiner le son sur la pellicule (voir note 21). Dobson défend l'idée que McLaren, à cette époque-là, ne connaissait pas ces expériences, et donc qu'il a découvert la technique par lui-même. Malheureusement le premier film dans lequel McLaren a essayé d'utiliser le son dessiné, *Book Bargain*, de 1937, a été finalisé avec une conception sonore différente, puis le chef de McLaren, Alberto Cavalcanti, pensait que ce type de sons artificiels entraient en conflit avec la voix de commentaire du film –il s'agissait d'un documentaire-.

En tout cas, McLaren travailla beaucoup avec cette technique. Ses premiers films faits à New York pour la collection de films abstraits du musée Guggenheim dans les années 40 ont été développés avec la technique du son dessiné sur la pellicule, les premiers étant *Dots* et *Loops* (1940). Dans ses propres paroles : « En travaillant constamment avec la pellicule dans la salle de montage, j'en vins à me rendre compte que la base du cinéma n'était rien d'autre qu'une bande de celluloïd dont la longueur représente le TEMPS. Ma première impulsion : dessiner directement sur le film sans me servir d'une caméra me parut de plus en plus légitime» [cité par (Bastiancich, 1997, p. 50)]. En 1951 McLaren participa à un documentaire sur cette technique du son animé, nommé *À la Pointe de la Plume*, réalisé par Don Peters et Lorne Batchelor. Dans les films produits de cette façon McLaren travaillait simultanément sur la bande image et sur la bande-son image par image, ce qui lui permettait d'accoupler des éléments visuels et des éléments sonores tout en réalisant le film. Alfio Bastiancich a interprété cette pratique créatrice d'une façon très intéressante pour notre recherche :

*La correspondance entre temps filmique et longueur de la bande de celluloïd est caractéristique d'une conception du cinéma libérée de tout conditionnement littéraire et narratif. Dans la technique du dessin direct McLaren utilise la pellicule comme un peintre se sert de la toile, avec pour résultat la production (en milieu non narratif et non représentatif) de ce que Gianfranco Bettetini définirait comme une « correspondance entre le temps de l'énonciation et le temps de l'énoncé » (BASTIANCICH, 1997, p. 50).*

On peut remettre cette idée dans le cadre conceptuel ici développé : en éliminant la médiation de la caméra et des microphones dans le dispositif matériel de production, l'animation directe va au-delà de l'animation traditionnelle, laquelle déjà avait supprimé la présence des comédiens, des scènes et des objets, mais non des voix et des sonorités, bien qu'ils aient été produits et couplés dans l'étape de postproduction. Le dispositif d'énonciation de l'animation directe se rapproche ainsi de celui de la peinture. Dans le cadre des pratiques

sémiotiques, on peut dire que le cinéma traditionnel coordonne une série de pratiques de production dans une stratégie générale, tandis que le cinéma d'animation directe réduit cette stratégie à une seule pratique créative, exécutée exclusivement par l'animateur<sup>176</sup>.

Le cas de la deuxième technique est similaire, mais avec une particularité qui affecte le niveau des signes : tandis que le « peintre sur pellicule » travaille avec des signes expressifs ouverts et singuliers (comme le peintre traditionnel) l'animateur du son qui travaille avec le système des cartons dessinés a un « code » plus ou moins stable. Le système de Voinov peut n'être pas exactement le même que celui Pfenninger, et celui de McLaren n'est certainement pas le même que celui de ses précurseurs, mais tous les trois travaillent avec le même principe, avec un répertoire fixe d'unités graphiques/sonores qui peuvent se combiner selon des règles de composition. Dans le cas de McLaren, c'est dans les années 50, en travaillant à L'Office National du Film du Canada (ONF) qu'il a commencé à utiliser son système de cartes, basé sur celui de Pfenninger et Voinov. McLaren le décrit ainsi :

*J'ai préparé de cette façon un grand nombre de cartons, chacun d'environ 40 cm de longueur sur 4 cm de largeur. Un carton pour chaque demi-ton de cinq octaves. Les cartons sont dans une boîte que je garde à côté de moi quand je fais les prises de vue, sous l'œil de la caméra d'animation. La boîte me sert, pour ainsi dire, de clavier. Je choisis ainsi un carton pour la fréquence dont j'ai besoin. Je la pose sur la table devant l'appareil et je la photographie toutes les fois que cela est nécessaire. Pour une note courte, je reprends seulement un photogramme. Pour une note longue, une vingtaine ou même une trentaine. Si le volume est piano, je recouvre une partie du dessin avec une feuille noire. Si, par exemple, c'est un crescendo, j'enlève progressivement la feuille pour découvrir le dessin – et ainsi de suite. (MCLAREN, NOTE SUR LA MUSIQUE SYNTHETIQUE, 1997, p. 162)*

On peut faire maintenant un parallèle entre les deux techniques du son dessiné. D'après Aime Mollaghan, il y a deux écoles de pensée autour du son animé. La première est

---

<sup>176</sup> Sur ce sujet, retenons la description suivante, faite par André Martin en 1958, de l'*ethos* créateur de notre animateur : « La solitude de McLaren devant ses films n'est pas seulement celle d'un artiste devant sa toile, d'un poète devant son texte, mais ressemble aussi à celle du savant aux prises avec les éléments matériels rebelles de sa recherche. Le mot préféré de ce cinéaste, dont les films se déroulent avec le naturel des associations d'idées, est 'méthode'. C'est avec le souci constant de la technique et de l'instrumentation qu'il avance par tâtonnements et essais successifs. Avec un sérieux scientifique, il assemble des matériaux, enregistre leurs actions, compare un grand nombre d'essais, inventant des styles inédits de composition » (Martin, 2000, p. 234).



représentée par Pfenninger et Moholy-Nagy, et le McLaren de *Synchromie*. Cette « école » est caractérisée comme étant « intrinsèquement logique, évitant un discours esthétique », son but principal étant le « développement technique qui a créé une nouvelle forme d'écriture sonore libre des contraintes provenant des institutions et de la notation existantes ». L'autre école est représentée par Fischinger et ses expériences avec le son animé. Dans cette école l'intérêt principal était « la création d'une unité plus grande entre l'image et le son », et en outre, « l'exploration des relations entre les formes graphiques et leurs homologues sonores ». L'héritage de cette école sera recueilli par Guy Sherwin (MOLLAGHAN, 2015, p. 98).

D'après nous, on peut encadrer ces « écoles de pensée » dans la perspective des niveaux de pertinence, en prenant comme centre du phénomène le niveau des pratiques et, particulièrement, de *l'acte créatif du dessin sonore en tant que pratique*. En fait, le dessin sonore est un cas exemplaire du parcours de la manifestation proposé par Fontanille. Pour décrire ce parcours, nous avons la chance d'avoir plusieurs matériaux descriptifs : textes, documentaires, interviews et témoignages décrivant la pratique de l'animation du son, les techniques, les supports, etc. Nous allons alors commencer l'approche de ce phénomène par des textes, verbaux et audiovisuels, mais non pas pour en faire une description sémiotique, mais comme une méthode de constitution des données d'analyse de la pratique de l'animation sonore. Nous sommes conscients que la description d'une pratique sémiotique devrait idéalement être « en acte » ; c'est-à-dire, à partir de l'observation directe du processus. Néanmoins, nous croyons qu'il est possible de faire une description sémiotique des pratiques du passé –dans une recherche historique, par exemple- à partir des matériaux et des documents dérivés ou périphériques au phénomène examiné. C'est-à-dire, en utilisant une méthode intertextuelle. Dans notre cas, nous disposons, à portée de main, de textes et de documentaires qui ont la plupart un accent fortement « métatextuel », que nous allons exploiter pour la description de cette pratique et de ses relations avec les autres niveaux. La pratique de l'animation sonore est donc décrite ici comme une pratique « type », c'est-à-dire, comme un ensemble d'opérations et de procédures plus ou moins standardisées, et non comme la pratique concrète de production d'un « texte filmique » spécifique.

Selon Gianni Rondolino, parmi les attitudes du travail de cet animateur, sont d'une grande importance : l'expérimentation, l'attitude ludique et l'attitude didactique (Rondolino, 1997, p. 13 - 26). Nous avons déjà cité deux textes de grand intérêt pour la description de cette pratique : (McLaren, *Son Animé sur Pellicule. Les Débuts*, 1997), de date inconnue ;

(McLaren, Note sur la musique Synthétique, 1997) originalement de 1951; et nous allons inclure dans son intégralité un autre texte de notre animateur : (McLaren, Quelques Notes sur le Son Animé Utilisé dans les Films : Now is the Time, A Phantasy, Voisins et Two Bagatelles, 1997), originalement de 1952. En plus de ces textes, il y a trois documentaires qui expliquent ces techniques, et qui, d'une certaine manière, situent et décrivent la scène pratique de l'animation sonore et ses actants : *À la pointe de la plume* (Don Peters et Lorne Batchelor, 1951) ; *Experimental film* (Arthur Lipsett, 1962) ; et *l'animateur musicien* (Éric Barbeau, 2004). Il y a aussi deux autres documentaires qui recueillent des entretiens avec l'animateur avec 14 ans d'intervalle : *The eyes Hears, the Ears Sees* (Gavin Millard, 1970) ; *McLaren on McLaren* (Grant Munro, 1983). Avec ces références et l'information fournie par les commentateurs de l'œuvre de McLaren, nous pouvons entreprendre cette description.

### **La scène pratique**

Dans les documentaires mentionnés, on trouve une première description de la scène pratique de l'animation sonore et de ses actants. Ainsi dans le film *l'animateur musicien* (Éric Barbeau, 2004), le réalisateur distingue trois techniques de son dessiné chez McLaren : 1. Peinture sur pellicule claire ; 2. Rayures sur pellicule noire ; 3. Sons photographiés sur pellicule. La différence entre les deux premières est celle du support (pellicule claire ou noire), et des instruments d'inscription sur le support (pinceaux et stylos pour la pellicule claire ; couteaux, ou lames de rasoir pour la pellicule noire). La différence entre ces deux techniques et la troisième est toutefois plus radicale, puis cette dernière comporte le dessin des sons sur autre support : les cartons, et une étape de plus: l'enregistrement photographique des dessins sur les cartons<sup>177</sup>.

---

<sup>177</sup> Pour cette raison, il est plus fréquent de distinguer seulement entre deux techniques : l'animation sonore directe et l'animation par cartes dessinées.

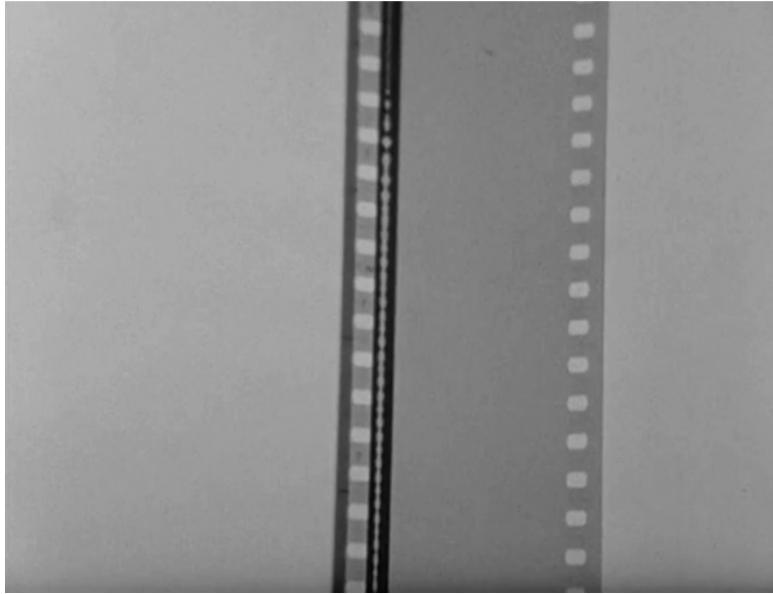


Figure 14 : Le support : pellicule. *À la pointe de la plume*  
Source : [Don PETERS et Lorne BATCHELOR, 1951]

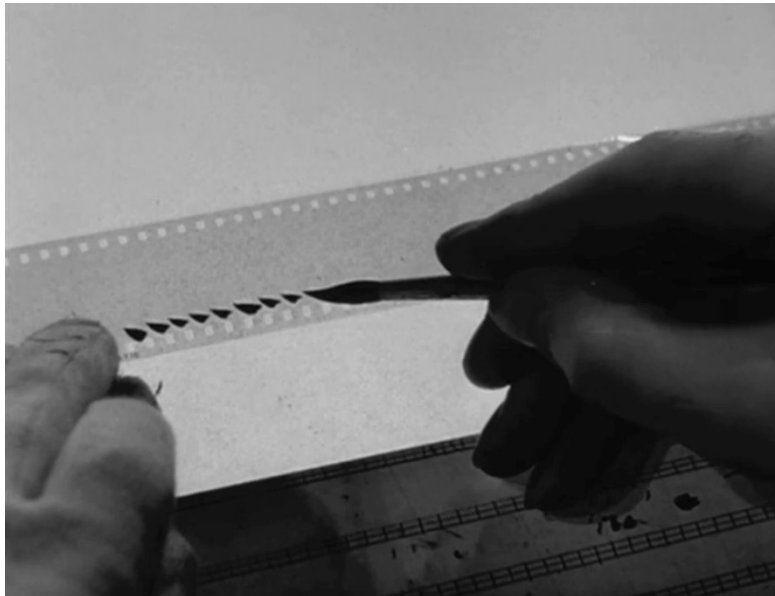


Figure 15 : L'opération : dessin sur pellicule claire. Photogramme du film *À la pointe de la plume*  
Source : [Don PETERS et Lorne BATCHELOR, 1951]

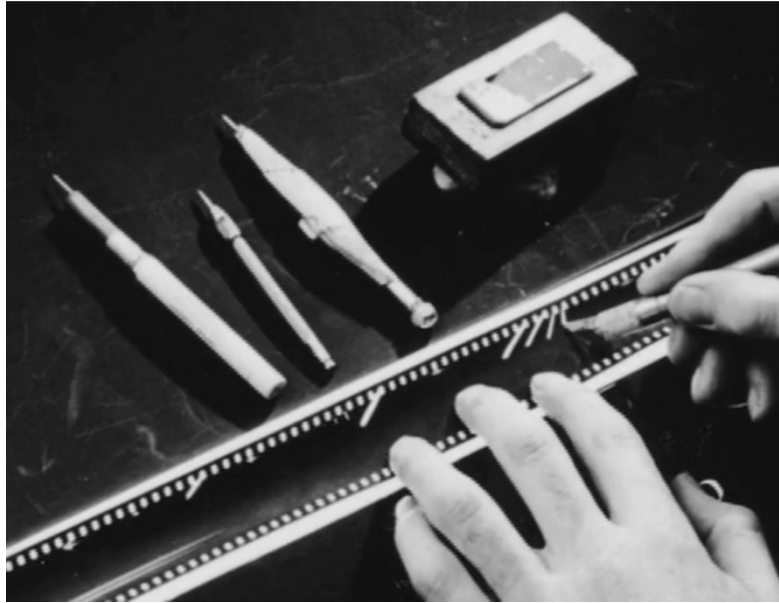


Figure 16 : L'opération : rayure sur pellicule noire. Photogramme du film *L'animateur musicien*

Source : [Éric BARBEAU, 2004]



Figure 17 : Les objets-adjuvants : banc et moviola. Photogramme du film *Experimental Animation*

Source : [Arthur LIPSETT, 1962]



Figure 18 : Les objets-adjuvants : la moviola (image-gauche, son-droite). Photogrammes du film *l'animateur musicien*

Source : [Éric BARBEAU, 2004]



Figure 19 : L'opérateur. Photogramme du film *À la pointe de la Plume*

Source : [Don PETERS et Lorne BATCHELOR, 1951]



Figure 20 : Les adjuvants : Grant Munro (Gauche) et Evelyn Lambart (droite). Photogramme du film *McLaren on McLaren*

Source : [Grant MUNRO, 1984]

L'actant-opérateur est l'animateur : McLaren lui-même ; il dessine, il strie, il coupe, il compose... Quelque fois, il y a des *adjuvants humains*, ses collaborateurs, comme Grant Munro et Evelyn Lambart<sup>178</sup>. À part l'objet-support et des objets-instruments d'inscription, il y a aussi des objets et appareils adjuvants : c'est le cas, principalement, de la *moviola* et du banc. Dans le film *Experimental Animation* (Arthur Lipsett, 1962), McLaren déclare : « Je passe la plupart de mon temps ici, devant cette machine et ce banc. Parce que, sans sortir le film de la machine, je peux le tirer par ici et je peux dessiner ou gratter des images et dès que j'en ai gratté quelques-unes, ou parfois même seulement une, je peux la faire passer » (voir figures 10 et 11).

Dans le cas de l'animation sonore par cartes dessinées on peut même parler de stratégie, en tant qu'il y a plusieurs pratiques coordonnées : le dessin des ondes sonores sur cartons ; la composition et combinaison syntaxique des unités graphiques/sonores ; le processus photographique : cadrage, caché ou coupage des stries sur les cartes, etc. (Voir ci-dessous le texte de McLaren décrivant cette technique). Dans ces pratiques il y a aussi des objets-adjuvants (la caméra, les manettes et la rainure du banc d'animation –Figure 16-) et des objets d'inscription (la boîte de cartes –Figures 14 et 15–). C'est pour cette technique que McLaren a bénéficié de la collaboration d'Evelyn Lambart.

<sup>178</sup> Il y a eu d'autres collaborateurs, mais Munro et Lambart ont travaillé avec McLaren spécifiquement dans des films avec du dessin sonore. En fait, Lambart peut être considérée co-créatrice de la technique d'animation sonore par cartes.



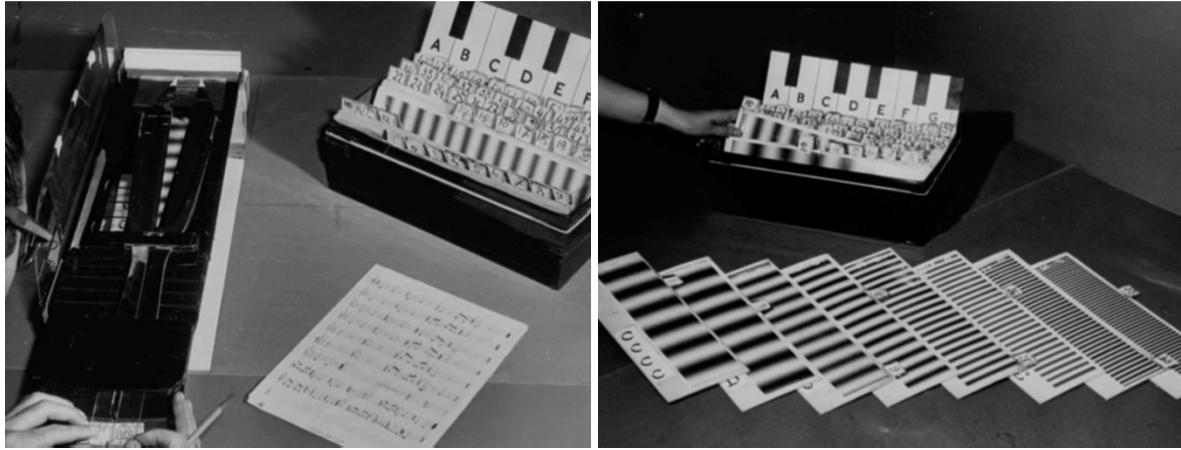


Figure 21 : Les objets d'inscription : les cartes. Photogramme du film *McLaren on McLaren*

Source : [Grant Munro, 1984]



Figure 22 : Les cartes. Photogramme du film *l'animateur musicien*

Source : [Éric Barbeau, 2004]





Figure 23 : Le banc d'animation (Mainettes et rainure). Photogramme du film *l'animateur musicien*

Source : [Éric Barbeau, 2004]

Dans les deux techniques, le dessin sonore est une pratique (ou un ensemble de pratiques) qui doit être accommodée avec d'autres pratiques dans une stratégie de création : premièrement, avec la composition de la bande-image, qui est aussi une autre pratique, bien que dans ce type d'animation, les deux peuvent arriver à se confondre, ou à s'échanger. Et après avec d'autres opérations de la production audiovisuelle. Bien que l'animation directe tende à éliminer ces autres opérations en concentrant tout travail créatif dans l'animateur, il y a parfois des pratiques de mixage du son ou de travail sur l'image, communes aux pratiques du cinéma en général.

Quant à la pratique créative du dessin sonore et de son articulation avec la composition visuelle, il s'agit surtout de connaître les correspondances entre le temps et l'espace et sa segmentation ; entre la taille, la forme et les distances –c'est-à-dire, les propriétés visuelles des traces et rayures- et les propriétés des sons qui en résultent : volume, ton, fréquence, etc. Cette correspondance temps/espace est particulièrement importante dans la composition musicale d'après le son dessiné ; comme l'a dit McLaren :

*Je marque au crayon gras la terminaison de chaque phrase musicale et l'emplacement correspond à la longueur de cette bande sonore. Alors que généralement le travail du compositeur de musique consiste à transformer, si l'on peut dire, des centimètres en secondes, le mien, tout au contraire, consiste à transformer des secondes en centimètres. Une note vibre pendant un temps donné : il me faut créer (au rythme de 24 images à la seconde) tant d'images pour couvrir cette note. [MCLAREN, cité dans (MARTIN, 2000, p. 245)]*

Dans le cas de l'animation sonore directe, McLaren a développé lui-même ces compétences, à partir de l'expérimentation. Il les présente avec un certain détail dans le film *À la pointe de la plume* (Don Peters et Lorne Batchelor, 1951). En ce qui concerne la technique des cartes, on a déjà vu que McLaren a dit qu'il a pris le « système » de Pfenninger, en l'adaptant avec Evelyn Lambart<sup>179</sup>. Il s'agit d'une procédure fortement codifiée, une pratique basée sur un transcodage visuel du son et de ses propriétés dans un système réglé d'unités clairement segmentées, ce qui nous mène au niveau des signes.

### **Les signes et les pratiques. Le clavier de papier de Norman McLaren**

Des techniques d'animation du son pratiquées par McLaren, la plus intéressante semble être l'animation directe : plus « expressive » et spontanée, avec une plus petite distance entre l'acte vivant d'énonciation et l'énoncé résultant, et la même pratique productrice pour le son et l'image : le dessin direct... On pourrait même dire que l'animation directe du son a un plus grand intérêt esthétique, tandis que l'animation du son dessiné sur cartes a un intérêt plus technique, même scientifique.

Néanmoins, cette dernière technique a été utilisée par McLaren de façons très versatiles et diverses : pour remplacer toute la bande sonore (musique, paroles, bruits) dans un film comme *Voisins*, fait avec des acteurs vivants et des scènes réelles selon la technique de la « pixilation » (capture image par image des objets et personnages) ; comme un autre « instrument », combiné avec les instruments traditionnels, comme dans *A phantasy*, où le son synthétique est combiné avec des saxophones ; ou simplement pour composer la musique en tant que registre exclusif de la bande-son. Quoi qu'il en soit, cette technique permet de produire une plus large gamme de sons et un degré de contrôle des traits du son plus que toute autre technique avant l'ère numérique.

À cet égard, les notes de McLaren et ses collaborateurs de l'Office National du Film de Canada sont un document essentiel pour comprendre cette pratique. Nous les incluons dans leur intégralité :

---

<sup>179</sup> Nous n'avons pas trouvé d'information précise sur la source où McLaren a trouvé la méthode de Pfenninger (C'est-à-dire, comment il a « acquis cette compétence »). On a déjà mentionné les documentaires que circulaient dans les années 30, et que McLaren connaissait. Néanmoins il semble que la description de cette technique n'était pas trop précise. En tout cas, non seulement il connaissait la technique, mais il l'a développée et standardisée.

## ***QUELQUES NOTES SUR LE SON ANIMÉ UTILISÉ DANS LES FILMS*** <sup>180</sup>:

### ***NOW IS THE TIME, A PHANTASY, VOISINS ET TWO BAGATELLES***<sup>181</sup>

*Il y a beaucoup de façon de réaliser le son animé, certaines ont été expérimentées depuis 1930. Ces notes et observations traitent seulement de la méthode particulière qui a été développée dans les dernières années à l'Office National du Film du Canada.*

### ***LES DESSINS UTILISÉS POUR LES ONDES SONORES***

*Comment sont les dessins des ondes sonores et comment les obtient-on? Il existe différentes façons de les faire. Par exemple, il aurait été possible de les faire en enregistrant (c'est-à-dire en photographiant) des sons musicaux réels sur la bande-son; et donc en copiant les formes résultantes de la bande. Cependant, cette méthode est aussi inutile et créativement nulle comme l'est la réalisation de dessins animés en photographiant des acteurs vivants et en traçant ensuite leurs contours.*

*Dans les films en question, une approche qui n'était pas naturaliste fut ensuite tentée, sans essayer d'imiter les sons naturels ou les instruments de musique traditionnels.*

*De nouveaux types d'ondes sonores furent réalisés, en utilisant des formes simples et faciles à dessiner.*

*Les dessins consistent en une simple forme de base qui est tout le temps répétée jusqu'à l'obtention d'une bande avec un motif. La figure peut aussi être seulement un segment répété qui court à angles droits dans la même direction que la bande-son sur un fond obscur; la répétition identique de la figure permet de donner lieu à une série d'ondes sonores ayant un timbre constant et défini.*

### ***TONALITÉ***

---

<sup>180</sup> Pour une description des films 3D de Norman McLaren, voir (LUCENA JÚNIOR, 2002, pp. 94-95).

<sup>181</sup> Norman McLaren, *Some Notes on Animated Sound as Developed at the National Film Board of Canada*, polycop. Interne, ONF Canada, 1952, révision 1961, 5 pp. Voir (McLaren, *Quelques Notes sur le Son Animé Utilisé dans les Films : Now is the Time, A Phantasy, Voisins et Two Bagatelles*, 1997).

*Chaque fiche de notre archive de sons porte une bande de formes répétées entre une aire large d'environ un pouce et longue d'environ douze pouces. Il y a une fiche pour chaque demi-ton de l'échelle chromatique, pour un ensemble de cinq octaves, des deux octaves sous le LA (440 Hz) aux trois octaves en dessus. Pour la note la plus basse, la forme est répétée de quatre à cinq fois et pour la plus haute, elle est répétée environ cent cinquante fois.*

*Le principe de base qui détermine le nombre de fois où le motif doit être répété est celui-ci : le LA est composé de 440 vibrations par seconde, c'est-à-dire que pour enregistrer le LA sur une pellicule, la forme doit être répétée 440 fois par seconde, ou en termes de photogramme, 440 divisé par 24 (pour un projecteur réglé sur 24 photogrammes par seconde), ce qui fait 18,33 fois. Ainsi la fiche pour le LA porte la forme répétée 18,33 fois.*

*Pour augmenter d'une octave n'importe quelle tonalité donnée, le nombre de vibrations par seconde doit être redoublé, de telle sorte que la fiche pour une octave supérieure au LA porte la forme répétée 36,66 fois (deux fois 18, 33). Pour une octave en plus de cette dernière 73,33 (deux fois 36, 66). Pour l'octave inférieure au LA, 9,165 (la moitié de 18,33).*

*La subdivision d'une octave est un problème plus complexe. Il existe de bons articles sur ce sujet dans le cadre de bonnes encyclopédies de musique et dans l'Encyclopédie britannique. Celui-ci est toutefois trop ample pour être affronté ici. Nos fiches de tonalité sont marquées de notations musicales normales et rangées de façon systématique dans un petit récipient, pour former une espèce d'échelle musicale.*

*Quand on enregistrait la musique sur la pellicule, la boîte était placée près de la caméra, de telle sorte que le compositeur (qui était le technicien de la caméra), s'il désirait une note particulière, pouvait sélectionner du récipient la fiche correspondante et la mettre devant l'objectif.*

*Lors de passages occasionnels de tonalités très profondes, la musique était reprise à l'octave supérieure et à une vitesse double de celle du résultat final souhaité ; durant le réenregistrement la vitesse était réduite de moitié, en diminuant ainsi d'une octave la tonalité.*

### **LA NATURE EN MOSAÏQUE DE LA MUSIQUE**

*La caméra filme les fiches par intermittence, par photogrammes (au lieu de fonctionner de façon continue, comme un appareil normal pour enregistrer le son) : la musique a une nature en mosaïque, en d'autres termes elle se forme à travers la somme de petites unités, chacune d'elles durant un vingt-quatrième de seconde.*

*Si l'on désire que la durée de la note soit plus longue, il suffit de faire plusieurs photogrammes de la même fiche à la suite, en atteignant ainsi un effet soutenu à travers la répétition très rapide de la même note, comme sur une mandoline ou un xylophone; pour les notes très courtes, un ou deux photogrammes sont suffisants. Durant les pauses et les intervalles, on photographie une fiche noire.*

## **DYNAMIQUES**

*Avant d'exposer la pellicule, le compositeur doit déterminer le volume précis, ou le niveau dynamique, de la note. C'est une des innovations les plus récentes et les plus intéressantes de la musique animée, puisque dans le passé, les niveaux dynamiques n'ont jamais trouvé une codification précise, même pas dans la musique traditionnelle. Les indications standard pp, p, mf, f, ff, etc. indiquent une quantité de volume approximative et complètement relative et ne sont jamais appliquées à toutes les notes de la partition, leur détermination finale est laissée à l'interprétation de l'artiste. En créant la musique animée au contraire, la dynamique précise de chaque note de la partition est du ressort du compositeur: en d'autres termes, le compositeur doit aussi être un interprète.*

*À cette fin, nous utilisons vingt-quatre degrés de niveau dynamique (qui représentent l'échelle des décibels) et devant chaque note de la partition, on écrivait le nombre du niveau dynamique souhaité. Par exemple, 0, 1 et 2 représentent trois degrés différents de ppp ; 9, 10 et 11 trois nuances différentes de mp; 12, 13 et 14 trois degrés de mf; 21, 22 et 23 trois degrés de fff et 24 représente ffff. Les subdivisions de ces vingt-quatre degrés étaient constamment utilisées (en particulier dans les crescendos et dans les diminuendos), mais elles étaient rarement reportées dans la composition. Dans les crescendos locaux ou rapides et dans les diminuendos, juste les notations dynamiques de début et de fin étaient reportées, et les différents types de crescendo et de diminuendo (comme "arithmétique" ou "géométrique") étaient indiqués par un petit signe.*

*Le volume était contrôlé en manipulant l'obturateur ou le diaphragme de la caméra, en influençant ainsi l'exposition (contrôle de la densité variable); mais plus souvent on couvrait le dessin, large d'environ un pouce, jusqu'à ce que seulement un demi-pouce ou un quart-de-pouce restent visibles (contrôle de l'aire variable). La mesure était toujours faite en décibels, sans tenir compte de la méthode utilisée.*

## **DONNER UNE FORME AU SON**

*Le compositeur n'avait pas seulement le dernier mot, décisif, sur la dynamique des sons, mais il était aussi contraint de spécifier le contour exact de chaque note : c'est-à-dire le type*

*d'introduction, de prolongement et de fin que chaque note devait avoir. Ceci est très important, puisque le contour qui détermine l'effet "instrumental" est plus important que la qualité sonore fondamentale de la note. Dans le cadre des sons musicaux traditionnels, par exemple, une note de piano a une introduction très rapide, pas de période de prolongement, mais une longue période de fin: son contour est comme le sommet d'une montagne, comme un côté très raide et l'autre en pente douce. Une note typique d'orgue a une introduction soudaine, un vaste prolongement et un épuisement (extinction) très rapide: un contour qui ressemble à une plaine, avec un précipice d'un côté et une descente très raide de l'autre. Une percussion sur un bout de bois a un début brusque, aucun prolongement et s'éteint très rapidement. Les instruments à vent ont une forme d'introduction bien moins soudaine que les instruments à percussion. Un violon, comme la voix humaine, est quasiment capable de n'importe quel type d'introduction, de prolongement et de fin.*

*Ainsi le compositeur, en fournissant un contour particulier à chacune des notes, influençait celle qui normalement serait appelée la qualité instrumentale. Ceci s'obtenait en plaçant des taches noires, de différentes formes, devant la fiche de ton choisi, laquelle portait le dessin des ondes sonores; de cette manière, nous obtenions environ six types de contours du son, certains impossibles à réaliser avec des instruments traditionnels.*

*Dans Voisins, nous faisons un usage considérable des contours de son variables, tandis que dans les autres films dominait un contour à percussion (de nature cunéiforme).*

### **QUALITÉ DE TIMBRE**

*Dans la bande-son de Voisins, le nombre et la variété, des effets sonores et des qualités de timbre furent amplifiés de façon considérable par l'utilisation de différents groupes supplémentaires de dessins dont certains sons se levaient et se baissaient pour les effets de port et de glissade. Certains dessins, même si d'une préparation et d'un aspect très simple, présentaient une structure d'ondes sonores très complexes, pleine de notes harmonieuses qui produisaient des qualités sonores très stridentes et dures.*

### **HARMONIE**

*Pour de nombreux passages musicaux simultanés, aussi bien en harmonie qu'en contrepoint, trois différentes méthodes furent utilisées. Des dessins différents furent superposés l'un sur l'autre en de nombreuses expositions, ou plutôt la musique était divisée sur la longueur en plusieurs bandes parallèles et les différents dessins sont photographiés, l'un à côté de l'autre, dans chaque bande. On pouvait encore filmer chaque partie musicale sur une pellicule séparée, mélangeant ensuite les différentes parties pendant les nouveaux enregistrements.*

## **QUALITÉ ACOUSTIQUE**

*Le son animé est généralement "sec", sans résonance ni échos. Pour obtenir plus de résonance et une qualité acoustique plus grande, deux méthodes furent utilisées. La première, qui concernait les notes spécifiques et les effets localisés ou momentanés, fut réalisée en photographiant la même note dans une rapide séquence de volumes décroissants (c'est-à-dire, le même dessin en mesures de plus en plus petites); cela permettait de simuler l'effet naturel des ondes sonores qui rebondissaient en avant et en arrière sur les parois d'un instrument, d'une caméra, d'une salle, d'une caverne. Le degré où chaque note particulière de la composition peut être placée dans un tel cadre acoustique est contrôlé, pendant le tournage, par le numéro et la nature des répliques décroissantes du dessin original de la note.*

*Note de Maurice Blackburn sur la méthode utilisée pour composer la musique de A Phantasy.*

*La composition pour trois saxophones et son animé fut écrite normalement, en considérant le son animé comme un instrument d'ensemble. L'unique différence était que les notes synthétiques, dans leur longueur et place, étaient considérées en unités de 1/24 de seconde ou, en d'autres termes, en groupes de photogrammes.*

*De la reprise du son synthétique on tirait une trace de base (« click-track ») afin de synchroniser les autres instruments, qui consistaient en trois saxophones (alto, soprano et ténor). Le même instrumentiste joua du saxophone, les uns après les autres, guidé par la trace de base, ce qui permettait un contrôle complet et autonome de chaque morceau dans le message final des quatre morceaux.*

Ce texte décrit le résultat des recherches et des expériences de manipulation du support et de la technique, et fait partie de la pratique artistique de McLaren, mais aussi de son intérêt dans l'enseignement et la communication de ses résultats. Ces « notes » peuvent fonctionner comme une sorte de « mode d'emploi », à travers lequel on peut acquérir les compétences nécessaires, bien que la complexité de la technique fait qu'on doit posséder une certaine connaissance musicale. Dans *Pratiques sémiotiques* Jacques Fontanille dit à propos des stratégies qu'il est possible de les transformer en textes : « ce sont alors des recettes de cuisine, des modes d'emploi, des notices de montage, qui fonctionnent alors, para rapport aux situations elles-mêmes, comme des métadiscours » (Fontanille, *Pratiques sémiotiques*, 2008, p. 29). Cette déclaration vaut également pour la métadescription d'une pratique. L'intérêt pédagogique de ce texte est évident, comme son caractère métadiscursif. Il ressemble une



petite « grammaire » de la composition sonore par cartes et, dans ce sens, rend compte d'une compétence, un « savoir-faire » du niveau de la modalité actualisante.

Ce « clavier de papier » (comme l'a nommé McLaren) offre tout un répertoire d'effets et possibilités de production des sons. On peut systématiser ce répertoire selon le trait sonore qui est produit ou manipulé, le type de codification visuelle que permet cette production/manipulation et le type d'unité et/ou d'effet concret produit.

Tableau 2: Répertoire d'équivalences pour la production des sons

Trait sonore	Codification visuelle	Unités/effets
Tonalité	Nombre de fois où le motif doit être répété	Demi-tons
Durée (mosaïque)	Nombre de photogrammes de la même fiche	Extension temporelle des notes
Dynamique (volume)	Densité/aire variable du dessin	Décibels
Forme (qualité instrumentale)	Taches noires devant la fiche de ton choisi (6 types)	Contour : « aspects » (introduction, prolongement, fin)
Qualité de timbre	Groupes supplémentaires de dessins	Effets spécifiques (« port », « glissade », etc.)
Harmonie	Superposition de dessins Bandes parallèles Pellicules séparées et mélangées	Nombre de « lignes musicales » (en harmonie ou contrepoint)
Qualité acoustique	Le même dessin en mesures décroissantes	Résonances, échos

Nous sommes confrontés à l'un de ces rares cas où une pratique donne lieu à un « langage », où la technologie et la sémiotique se croisent. Car, bien que cette pratique ne puisse être que le "transcodage" du son en image (à cause de l'utilisation un peu conventionnelle de la technique pour la composition musicale traditionnelle chez McLaren et ces prédécesseurs), il est possible qu'elle ouvre un univers expressif complètement nouveau. Il est donc possible d'affirmer avec Robert Beyer, qui était fasciné par les possibilités du son sur pellicule à la fin des années 20 : « Avec ces autres méthodes, il sera possible d'atomiser le

son, pour construire à partir de ses éléments de base de nouvelles couleurs de tonalité, pour traverser un domaine de timbres d'une immensité presque cosmique, qui dépasse de loin les limites connues »<sup>182</sup>.

Ces notes, écrites en 1952, se rapportent à la production sonore des films *Now is the Time*, *A Phantasy*, *Voisins Et Two Bagatelles*, réalisés au cours des années quarante et au début des années cinquante. La technique a été utilisée dans d'autres films au cours des décennies suivantes (comme *Rythmic*, de 1956, ou *Mosaic*, de 1965). Mais le film définitif fait avec la technique des cartes de son, par sa complexité et extension, et aussi parce qu'il est le dernier, est sans doute *Synchromie*, de 1971. Ce film présente un attrait particulier : l'équivalence presque complète entre la bande image et la bande-son, à travers l'inclusion de l'une dans l'autre. L'analyse de ce texte audiovisuel nous permettra donc d'observer une manifestation concrète de cette pratique, avec un autre ingrédient : la possibilité de regarder les « signes sonores », et d'écouter les signes visuels simultanément.

### La recherche de la « synchronie » parfaite

S'il est vrai que la particularité de la *Synchromie* est l'inclusion de la bande-son sur la bande image, on ne peut pas oublier que ce recours était déjà apparu chez les pionniers du son animé. On a mentionné le fait que Voinov a utilisé cette application de la technique dans ses films, dont on peut regarder les « peignes » de papier qui produisaient les sons (voir Figure 17).

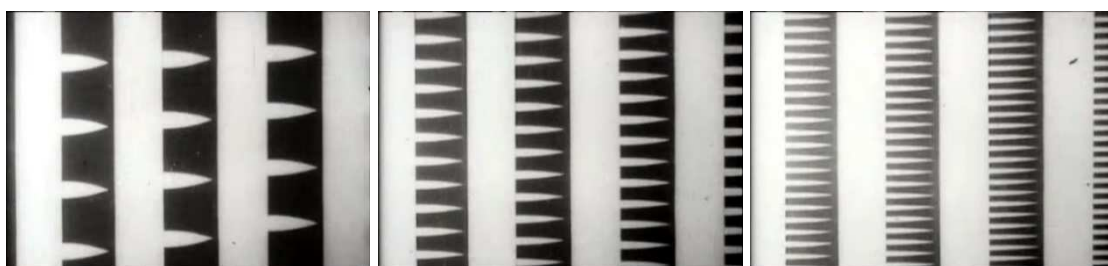


Figure 24 : Photogrammes du film *Sons sur papier*

Source : [Звуки из бумаги, 1932]

<sup>182</sup> “With these and other methods it will be possible to atomize sound, to construct from its basic elements new tone colors, to traverse a timbral domain of almost cosmic vastness, which far exceeds the known boundaries”. Cité dans (Patteson, 2015, p. 102).

Ces « peignes » ne sont pas visibles normalement, mais sont photographiés seulement sur la bande-son, pour qu'ils soient « lus » par le projecteur comme des sons. Voinov a inclus ces photographies aussi à l'écran comme un recours didactique, pour que le public puisse comprendre comment les variations de taille des pointes des « peignes » et de leurs distances produisent des variations de tonalité.

Pfenninger aussi, dans les *Tönende Handschrift* inclut l'image de la bande-son à côté de l'écran, et même la main de l'animateur sonore qui dessine les sons « en acte ». Bien sûr, il s'agit d'une simulation de cet acte, un cas d'« énonciation énoncée » plus proche de la figure de la métalepse, très fréquent dans l'animation des premiers temps, où la représentation de la main de l'animateur sur l'écran était fréquente <sup>183</sup> (voir Figure 18).



Figure 25 : Photogrammes du film : *Tönende Handschrift II*

Source : [Rudolf Pfenninger, 1932]

Dans ce film, entre le 0:56s et le 1:23s on peut voir la main de l'animateur dessinant les courbes du son qui sont censées être entendues. Après le son synthétique continue tandis qu'on cesse de regarder la bande-son, mais la main de l'animateur est toujours à l'écran, maintenant en dessinant des personnages qui interagissent avec elle : deux poissons. Cette séquence semble simple, mais son but est de montrer que la même main qui donne la vie aux caractères donne lieu à la musique qu'on écoute.

Le cas de McLaren est plus proche de celui de Voinov. Il ne veut pas montrer sa présence dans le procès, mais il montre l'artifice technique qui lui donne lieu : les cartes et ses dessins. Dans les premiers tests pour *Synchromie*, on peut constater la présence d'une seule bande-son à l'écran (Figures 19 et 20).

---

<sup>183</sup> A cet égard, voir par exemple (Limoges, 2012, pp. 132 - 133). Cet article se trouve aussi on-line dans <http://www.cinergie.it/?p=292>

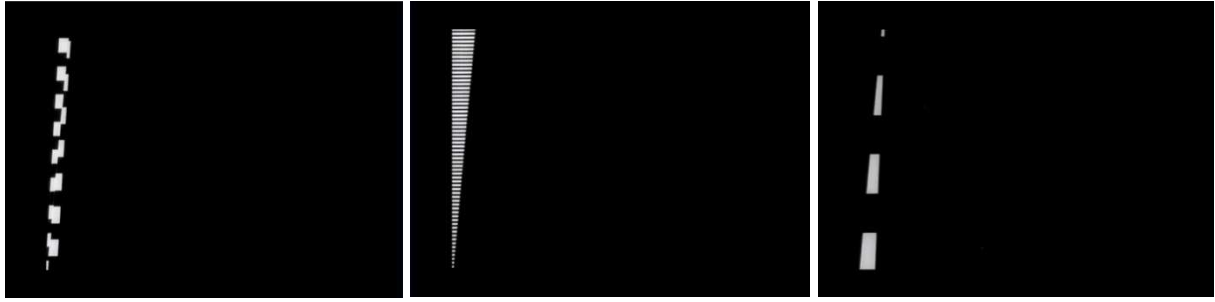


Figure 26 : Photogrammes du film : *Test A pour Synchronie* (fin des années soixante)

Source : [Norman McLaren, 196 ?]

Et l'image est plus proche de cette de *Synchronie* dans le test B :

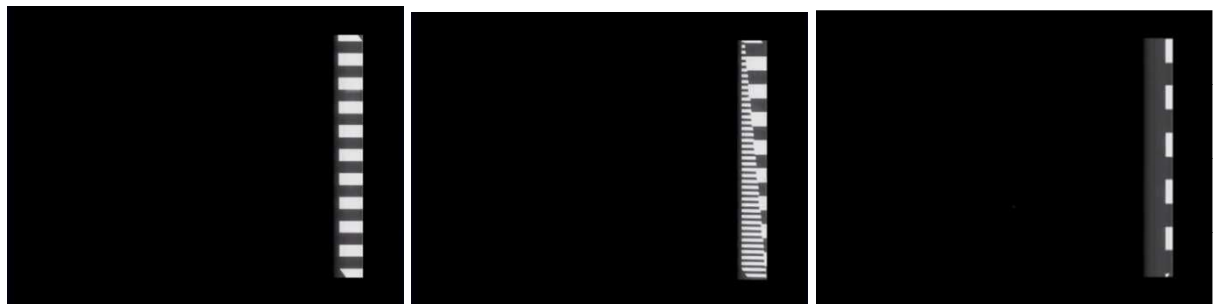


Figure 27 : Photogrammes du film : *Test B pour Synchronie* (fin des années soixante)

Source : [Norman McLaren, 196 ?]

La grande différence avec ces premiers tests est que dans le film définitif McLaren a introduit des modifications sur la bande-image, en multipliant les barres le long de l'écran. Ainsi McLaren prend à l'extrême le recours de montrer la bande-son. Le résultat : 7 minutes et 27 secondes de stimulation audiovisuelle dans sa forme la plus pure.

Alfio Bastiancich décrit ainsi ce qui se passe dans *Synchronie* :

*Les cartons furent organisés de façon à créer un thème musical de boogie-woogie, puis ils furent photographiés aussi bien sur l'espace réservé à la bande-son que sur celui de la bande image. Sur cette dernière, on arrive à avoir en même temps un maximum de onze cartons. Les couleurs, choisies arbitrairement, mais adaptées à la nature de la musique, furent obtenues à travers un système de filtres. Un tel procédé génère chez le spectateur l'effet d'un glissement de la notation musicale synthétique de l'espace qui lui est réservé à celui du champ visuel.*

*La composition de l'image est semblable à un « collage » géométrique où les signes, des couleurs très vives et contrastées, sont l'exposition visuelle des valeurs des intervalles*

données par le contrepoint sonore. La rigueur de l'orchestration figurative évoque de nouveau, comme déjà dans *Mosaïque*, les compositions picturales de l'Op-Art dont nous retrouvons ici les vibrations de couleur et l'instabilité de l'ensemble : ces deux phénomènes liés à la persistance rétinienne de l'image. (BASTIANCICH, 1997, pp. 105 - 106)

On a dit que l'équivalence est « presque » parfaite. Ce « presque » est dû à plusieurs variations arbitraires introduites par l'auteur, bien que dans le but de créer davantage d'autres correspondances avec le son. Ainsi, dans la description de Bastiancich on trouve des expressions comme : « Les couleurs, choisies arbitrairement, mais adaptées à la nature de la musique... ». La répétition même des cartons est arbitraire, parce qu'en réalité seulement une de ces cartes représente la bande-son.

D'autres auteurs ont essayé de montrer que toutes les variations visuelles introduites par McLaren ont des correspondances avec des traits sonores des autres ordres. Utako Kurihara a même fait la partition de la musique à partir de la dictée musicale de la bande-son, pour démontrer comment les variations visuelles dans *Synchromie* réussissent à « paraphraser en images » des aspects comme des hauteurs et registres, mais aussi « des lignes thématiques, la dynamique, des modulations, des séquences tonales, des agrandissements de motifs, le processus d'approche d'un couple de thèmes, des sensations contorsionnées en polyrythmie et l'effet de réverbération » (KURIHARA, 2005, p. 1).

Un autre auteur qui a essayé de montrer que les variations visuelles de *Synchromie* produisent d'autres équivalences avec le son est Terence Dobson, qui a fait ce qui est peut-être l'analyse la plus détaillée et systématique de ce film. Cet auteur identifie les modifications introduites par McLaren, dans l'idée que, en effet, chaque variation a une correspondance qui renforce l'effet de l'équivalence générale entre l'image et le son. Ainsi la multiplication des bandes, les couleurs appliquées et les mouvements latéraux au long de l'écran, ajoutés après la troisième partie du film, tous ces éléments coïncident avec les moments où la musique devient plus complexe (DOBSON, 2006, pp. 209 - 215). Dobson décrit aussi comment la musique du film est composée de trois thèmes musicaux, chacun enregistré dans une bande-son différente et puis mixés dans une seule bande. Ces trois thèmes ne sont pas différenciés à l'écran jusqu'à la troisième section du film, où chaque thème a son propre air visuel, en permettant une autre complexité en harmonie avec la musique.

Pour encadrer l'interprétation de *Synchromie* (comme le meilleur exemple du dispositif du son animé) dans la perspective des pratiques sémiotiques, nous faisons un

parallèle avec la lecture de ce film proposée par Dobson, dans le but de montrer que, au moins dans ce cas, on ne peut pas rendre compte de sa textualité sans passer à d'autres niveaux de pertinence. Le Dobson l'exprime lui-même avec ces mots : « La connaissance du procès de fabrication de *Synchromie* est nécessaire pour une compréhension plus complète du film » (DOBSON, 2006, p. 208)<sup>184</sup>.

### Une lecture structurelle de *Synchromie*

Nous voudrions, premièrement, identifier la structure générale du film, avec l'aide de la lecture de Dobson, mais aussi à partir de notre propre expérience d'audiovision. Cela nous permettra d'identifier les différents types de relations entre l'image et le son. Nous essayerons d'identifier des glissements de terrain entre les niveaux de pertinence, inévitables, d'après nous, pour rendre compte de ce type de phénomènes sémiotiques.

### Une structure « didactique » ?

D'après Dobson, il y a trois parties clairement différenciées dans *Synchromie*, séparées par deux « codas » (Voir Figure 21). La première partie va du début du film au premier fondu au noir, aux deux minutes et quatre secondes de la durée totale. Cette première partie fonctionne comme introduction et présente les conventions d'équivalence son/image, notamment celles relatives à la tonalité et le volume :

*Le ton est représenté visuellement dans le film par des striations horizontales à l'intérieur d'une bande verticale en fréquences différentes. Des stries étroites rapprochées équivalent aux sons aigus, tandis que des stries larges et espacées équivalent aux sons graves. Un tel système de représentation visuelle des sons (musicaux) est inhabituel (...) McLaren a réalisé cela. Il a commencé le film avec une démonstration des configurations du ton. La représentation du volume sonore est aussi clairement démontrée à ce stade. La relation des tonalités du son avec les fréquences visuelles affichées sur les bandes, par conséquent, peut être reconnue dans la première partie du film, la plus simple. La stricte conformité temporelle, note par note, aide le spectateur à identifier une relation auditive/visuelle. (DOBSON, 2006, pp. 215 - 217)<sup>185</sup>*

---

<sup>184</sup> "A knowledge of the making of *Synchromy* is necessary for a fuller understanding of the film".

<sup>185</sup> "*Pitch is visually represented in Synchromy by horizontal striations, within a vertical band, of differing frequencies. Narrow striations close together equal high sounds, broad striations further apart equal low sounds. Such a system of visually representing (musical) sounds is unusual (...) McLaren realized this. He began Synchromy with a demonstration (exposition) of the pitch*

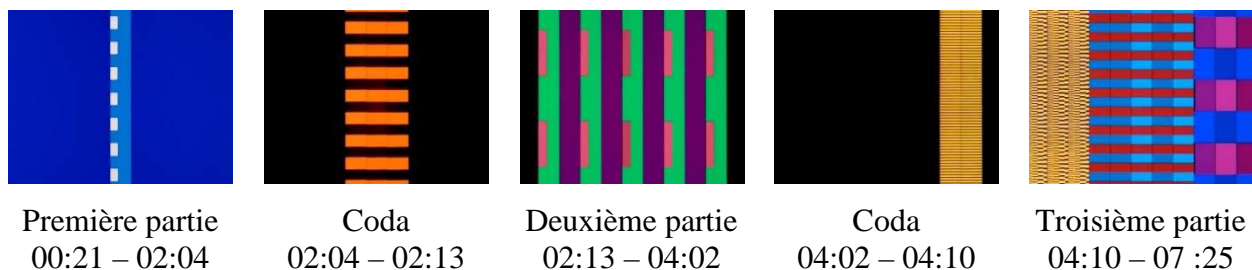


Figure 28 : La Structure générale du *Synchromie* et des photogrammes correspondants à chaque partie du film

Source : [Norman McLaren, 1971]

Après cette démonstration, la première « coda » commence et dure environ dix secondes (jusqu'à 02:13). Cet interlude présente la même correspondance entre le ton et la longueur et la distance des rectangles qui conforment les stries (voir Figure 24). Pourtant, pour fonctionner comme « ponctuation » du film, il y a une variation visuelle importante : le fond de l'image est noir et les stries de trois bandes sont jointes pour former une strie plus longue. Dans sa dimension sonore, cette coda présente des sons entrecoupés de silences.

La deuxième partie va jusqu'aux quatre minutes et deux secondes environ, avec un deuxième fondu au noir, où la seconde « coda » commence. Cette partie est caractérisée par la multiplication progressive des bandes colorées, en commençant par trois jusqu'à onze (en couvrant l'écran) ; pour l'accélération générale du rythme, l'élévation des tons et le soutien des aigus.

---

*configurations. The representation of sound volume is also clearly shown at this stage. The relationship of sound-pitch to the visual frequencies displayed on the bands, therefore, can be recognized in the earlier, simpler part of the film. The strict temporal accordance, note by note, helps the viewer identify an auditory-visual relationship”.*



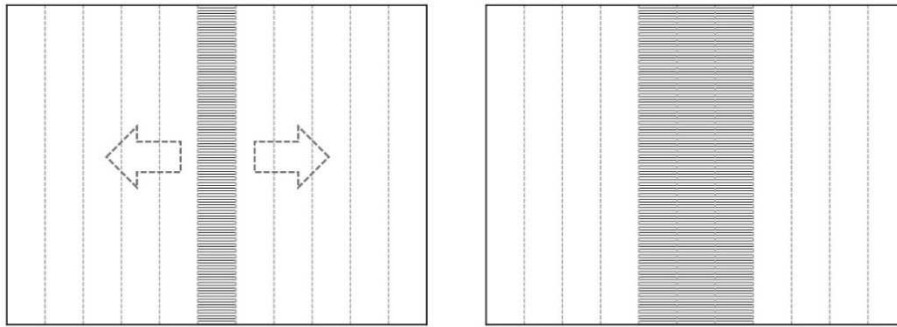


Figure 29 : Expansion des barres (première partie)

Source : [Image de l'auteur]

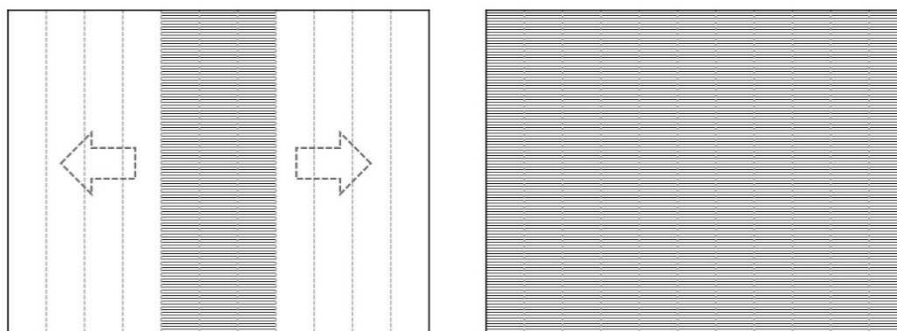


Figure 30 : Expansion des barres (deuxième partie)

Source : [Image de l'auteur]

La seconde coda dure huit secondes environ et, contrairement à la première, ici les stries se composent de deux bandes (voir Figure 24). Mais il y a encore une différence plus importante : d'après Dobson cette coda « montre une interprétation spatiale de la tonalité », puis

*...dans ce segment cependant, le ton est également indiqué par la position des bandes. Ces bandes larges apparaissent sur la gauche de l'écran si elles représentent un son faible, et à travers, et jusqu'à la droite si un son haut est dénoté. Des positions intermédiaires se produisent également. Bien que ce type de présentation de la hauteur ne soit qu'une petite partie de Synchronie, la coda s'adapte cependant à la structure globale du film (...) La coda, alors, représente encore une autre modification de la présentation visuelle de la piste sonore. L'interprétation spatiale de la tonalité apporte un plus grand accord entre les aspects visuels et auditifs du film. Ses fonctions structurelles prêtent également à l'aspect visuel du film un*

*peu de cohérence croissante, donnant un plus grand accord avec son homologue auditif.*  
(DOBSON, 2006, p. 218)<sup>186</sup>

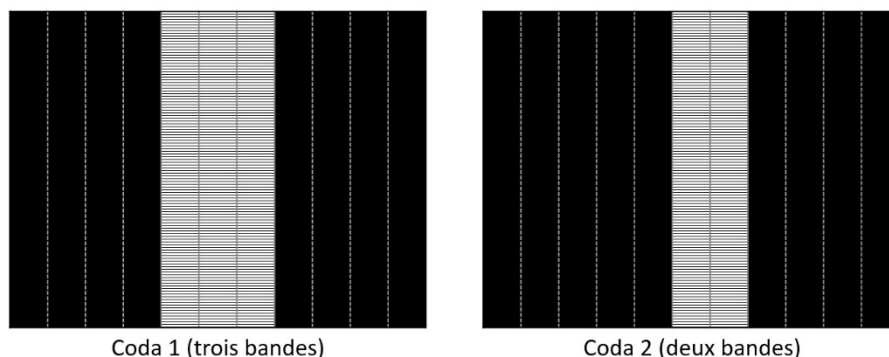


Figure 31

Source : [Image de l'auteur]

D'après Dobson, la deuxième coda est également une introduction effective à la troisième partie du film, puis le mouvement latéral des bandes anticipe les mouvements présents dans la section finale. Effectivement, cette troisième partie commence avec des mouvements latéraux des striations d'une bande à la suivante jusqu'à traverser l'écran, et imitant les déroulements des phrases musicales. Particulièrement, en conformité avec le déplacement latéral des stries dans la bande-son, il y a deux «voix» qui s'imitent rythmiquement l'une à l'autre, note par note. Les striations qui représentent chaque voix (distinguées aussi l'une de l'autre par la couleur : blanche et brune) sont placées initialement sur côtés opposés de l'écran, mais elles commencent après à se déplacer, bande par bande, vers le centre de l'écran, tandis que les deux lignes sonores alternent.

En plus de cette complexité rythmique, McLaren introduit une autre complexité visuelle: l'extension de la taille des striations à plus d'une bande, qui produisent des dissymétries visuelles entre les stries avec un grand impact sur l'écran. Ainsi, l'animateur parvient à produire des « macro-formes » qui renforcent les jeux entre les deux voix. C'est à ce point que la bande-son devient encore plus complexe, par l'apparition de trois pistes

---

<sup>186</sup> *“In this segment though, pitch is also indicated by the position of the bands. These wide bands occur on the left of the screen if they represent a low sound, and across to the right if a high sound is denoted. Intermediate positioning also occurs. Although this type of pitch presentation is just a small part of the film, the coda does nevertheless, fit the film's overall structure”.*

différentes, et de leur représentation en divisant l'écran en trois parties différentes. Avec la démultiplication des stries et de ces dissymétries, *Synchromie* atteint son apogée dans cette troisième partie, avec les trois lignes musicales en vitesses croissantes, jusqu'à donner lieu à un ralentissement représenté dans l'image par la disparition progressive des bandes. Ainsi le film fini comme il a commencé, avec une seule bande sur l'écran et, finalement, un fond noir et une note soutenue et le silence.

Terence Dobson souligne deux aspects intéressants de cette dernière partie du film. Le premier est le fait que bien que la synchronisation image-son soit en grande part construite avec des conventions internes du film et pas seulement le résultat de « montrer » le dessin visuel de la bande-son sur l'écran, les manipulations arbitraires introduites par McLaren se traduisent par un effet de synchronisation encore plus fort. Le deuxième aspect concerne les mouvements des striations et des barres et sa relation avec la nature des sons individuels :

*Le type dominant de mouvement dans le film est un scintillement hypnotique. Cela résulte de l'unité temporelle absolue entre le son et la vision. Dans *Synchromie* les sons, ou des notes, sont staccato et ont des silences très courts entre eux. Dans la dimension auditive, la continuité –la forme globale ou la mélodie des sons– demeure primordiale. Cela n'est pas le cas avec la présentation visuelle. L'effet de pulsation, qui est perçu comme à peine une caractéristique de la présentation du son, domine complètement la perception visuelle du film dans la mesure où la reconnaissance des expressions temporelles est sérieusement compromise. Pourtant, sur la piste sonore, l'organisation temporelle dans des schémas ou des phrases, malgré le staccato de la musique, est l'impression écrasante. (DOBSON, 2006, pp. 218-219)<sup>187</sup>*

D'après Dobson, bien que McLaren aurait pu utiliser des notes en *legato*, il a choisi ne pas le faire, parce que « l'effet de scintillement l'excitait au point qu'il a préféré ne pas le sacrifier pour obtenir une plus grande continuité visuelle » (DOBSON, 2006, p. 219).

---

<sup>187</sup> *“The dominant type of movement in the film is a mesmeric flicker. This results from the absolute temporal unity between sound and vision. In the film the sounds, or notes, are staccato and have very short silences between them. Aurally, the continuity -or overall shape or melody of the sounds- remains paramount. This is not the case with the visual presentation. The pulsating effect, which is perceived as just one feature of the sound presentation, completely dominates visual perception of the film to the extent that the recognition of temporal phrases is seriously impaired. Yet, on the sound track, the temporal organization into patterns or phrases, despite the staccato of the music, is the overwhelming impression”.*

Néanmoins, ce choix se traduit par une tension ou un conflit entre la perception continue de la musique (malgré le *staccato*) et la perception discontinue de l'image (le scintillement), à cause d'une caractéristique de ce qu'on a appelé le dispositif de réception :

*Dans le projecteur filmique le son est produit en déplaçant continuellement la piste sonore, avec ses stries, à travers la cellule photoélectrique. Du point de vue de l'« œil » –la cellule photoélectrique– les signaux viennent continuellement à un rythme très rapide. Une impulsion vélocité se produit. Certaines striations, quand elles sont soutenues, produisent un ton particulier. L'image visuelle dans le film, cependant, est produite par une séquence d'images statiques. Si la note est maintenue pendant plusieurs secondes, l'image sur l'écran sera statique. Ainsi le projecteur décode les images similaires des pistes visuelles et sonores différemment.* (DOBSON, 2006, pp. 219 - 220)<sup>188</sup>

Et même si la décodification des bandes visuelle et sonore était équivalente, ajoute Dobson, il n'y a aucune garantie que ces manifestations induisent chez le spectateur des états identiques, de sorte que l'effet final de ce film est l'alternance entre la confluence et le conflit des sons et des images.

Il est possible de modéliser ces relations et processus entre la pratique de création dont le produit est *Synchromie*, et sa réception, comme des programmes narratifs polémiques. Le premier programme concerne la pratique de production audiovisuelle du court métrage<sup>189</sup>. Ici, l'animateur est le sujet de faire (S<sub>1</sub>) qui cherche mettre en conjonction le sujet d'état, le spectateur (S<sub>2</sub>) avec un objet de valeur (Ov) : le film, et, particulièrement, la synchronisation entre les sons (et leurs traits) et les images (et leurs traits). Il s'agit d'un « faire-croire » ou, plus précisément, d'un « faire-percevoir » :

---

<sup>188</sup> *“In the movie projector sound is produced by continuously moving the sound track, with its striations, across the photo electric cell. From the point of view of the 'eye' -the photo electric cell- the signals come continuously at a very fast rate. A very fast pulse occurs. Certain striations, when sustained, produce a particular pitch. The visual image, in the film, however, is produced by a sequence of static images. If the note is held for several seconds, the image on the screen will be a static one. Thus the projector decodes the similar images of the visual and sound tracks differently”.*

<sup>189</sup> En ce sens, nous prenons la notion de production telle qu'est présentée dans le dictionnaire : « en sémiotique, la production est l'activité sémiotique, considérée comme un tout, et qui, située dans l'instance de l'énonciation, aboutit à la formation de l'énoncé » (GREIMAS & COURTES, 1993).

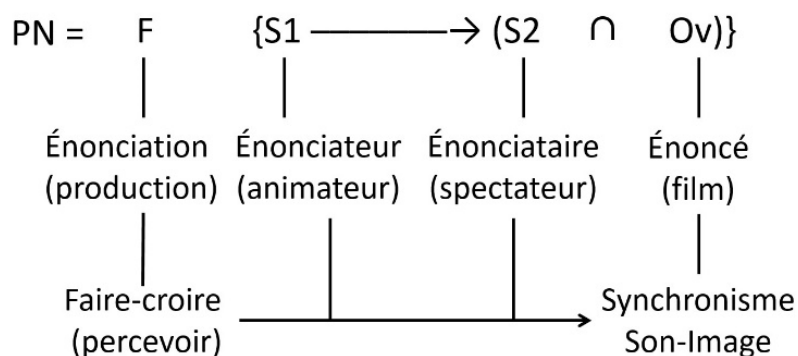


Figure 32

Source : [Image de l'auteur]

La structure même du texte-énoncé semble être conçue pour cette fin : la première partie et son caractère « didactique » peut-être lue comme la phase d'acquisition d'une compétence, la capacité de saisir l'équivalence entre les formants visuels et leurs correspondantes sonores. D'où sa simplicité visuelle et sa vitesse réduite : le spectateur doit « apprendre » les équivalences avant d'affronter des segments plus complexes.

Pourtant, cet apprentissage ne se limite pas à la première partie, mais il continue dans les autres, y compris les codas, particulièrement la seconde, où l'on doit apprendre la relation entre la position spatiale des barres et la tonalité. Dans cette coda on doit assimiler aussi le rôle d'un nouvel élément : le mouvement des barres. Néanmoins, à notre avis il ne s'agit pas encore du mouvement : les barres doubles apparaissent et disparaissent dans différents endroits de l'écran. Il s'agit donc d'un mouvement « potentiel » qui ne devient effectif que dans la troisième partie.

Dans la seconde, pour l'instant, il s'agit aussi d'« éprouver » la compétence acquise en face de l'accélération de la musique et la multiplication des barres. Dans cette partie la piste est doublée (et même triplée) pour de brefs instants, en annonçant la multiplication sonore de la troisième. L'effet global de cette partie est, pourtant, d'avoir « compris » le synchronisme et d'être en mesure de faire face à une plus grande complexité.

Certainement, dans la troisième partie la complexité atteint son point culminant. Les barres sont disposées en groupes selon leur conformité avec les trois pistes sonores, les couleurs sont rassemblées en trames et ces groupements visuels (formels et chromatiques) obéissent à la recherche d'équivalences avec les trois pistes sonores. Cette complexité

croissante de la bande-image et de la bande sonore produit un effet global de synchronisation totale.

Jusqu'ici, on peut rester au niveau de pertinence du « texte-énoncé » et parler d'un « énonciataire énoncé » qui correspond à l'ensemble des opérations décrites avant. Mais c'est à ce point que la description faite par Dobson nous invite à en sortir. D'une part, parce que seulement celui qui sait que les barres et les cadres qu'on voit à l'écran font partie de l'objet-support –la bande-son de la pellicule– et qui est montré comme une façon d'aboutir à la synchronisation, est capable de prendre conscience des « asymétries » ou des équivalences arbitraires décrites par Dobson : la multiplication de la bande-son en onze barres, les groupements, la spatialisation de la tonalité, les correspondances chromatiques<sup>190</sup>, etc. On doit mentionner que Kurihara identifie aussi des correspondances entre les formes visuelles et les configurations proprement musicales de la bande-son : l'expansion des « motifs » hors « thèmes » musicaux, représentée avec l'expansion des barres (à travers de groupements chromatiques) ; la modulation de la musique dans des intervalles courts de séquences harmoniques, représentée avec des déplacements de gauche à droite (et l'inverse) ou de changements des structures visuelles ; la représentation de la polyrythmie à partir de la séparation des arrangements de barres...<sup>191</sup> (Kurihara, 2005, pp. 3 - 5). En outre l'impossibilité technique (décrite par Dobson) de faire correspondre le flux sonore et le flux visuel dans la projection, implique la conscience des particularités du support de la scène pratique de la réception du film.

On peut distinguer les traits visuels et sonores équivalents dans l'objet-support, de ceux qui sont construits arbitrairement dans la postproduction du film :

---

<sup>190</sup> A cet égard, il est nécessaire de souligner que la colorisation du film est postérieure à la réalisation de la bande-son (et de la bande-image première ; le film est, originalement, en blanc et noir) donc toute « équivalence » ou « correspondance » entre la couleur et le son est tout à fait arbitraire.

<sup>191</sup> Nous ne fouillons pas cet aspect, tout d'abord, parce que nous voulons souligner les aspects proprement sonores (et non musicaux) du travail créatif de McLaren. Et deuxièmement, parce que nous ne disposons pas suffisamment de connaissances musicales pour un jugement réfléchi. À cet égard, nous nous référons au travail de Kurihara, quoiqu'incomplètement.

Tableau 3

Correspondances « motivées »	Correspondances « arbitraires »
Fréquence = longueur	Multiplication des barres = singularité de la piste sonore
Dynamique = largueur	Tonalité = couleur
Tonalité = nombre de motifs	Position spatiale = Tonalité
Qualité de timbre = groupes de dessins	Déplacements = séquences harmoniques
Harmonie = superposition des dessins	Staccato = scintillement visuel <sup>192</sup>
Qualité acoustique = mesures décroissantes des dessins	Groupement des barres (et des stries) = polyrythmie

On peut toujours, bien sûr, faire une lecture –ou une expérience perceptive– « immanente », en suivant les opérations inscrites dans le texte-énoncé audiovisuel, sans plus d'informations sur le « comment ça a été fait ». Une lecture qui suppose une syncope telle que décrite par Fontanille, en parlant des textes électroniques : « La « dématérialisation » du support des écritures, qui supprime le niveau de l'objet, nous fait directement passer du texte à la pratique » (Fontanille, *Pratiques sémiotiques*, 2008, p. 59). De la même manière, dans le cinéma le spectateur est confronté à une projection, et il n'a pas d'accès à l'objet-support : la pellicule. Seulement celui qui sait que ce qui est montré à l'écran est une représentation photographique de la pellicule peut capturer les variations et asymétries entre la bande-image et la bande-son, mais aussi découvrir les solutions créatives et la récursivité du réalisateur à l'heure de produire des interprétations visuelles du son et de la musique. C'est-à-dire, ce spectateur « conscient » du dispositif peut profiter non seulement de la richesse du texte, mais aussi de la pratique créatrice que le film offre aussi aux yeux et aux oreilles plus attentifs.

En ce sens, ce spectateur devient un « anti-énonciataire ». D'où le titre de Dobson : « confluence et conflit dans Synchronie ». Mais aussi le « paradoxe » identifié par cet auteur : c'est dans le conflit –l'identification des asymétries et des équivalences arbitraires– que le spectateur arrive à la jouissance complète du film.

<sup>192</sup> En fait, ce cas est particulier parce que le staccato est effectivement l'équivalent sonore du scintillement visuel, mais selon la lecture de Dobson, les particularités du dispositif de projection produisent chez le spectateur une perception lisse et continue de la bande-son, mais une perception intermittente de la bande-image.



On peut penser aussi que cette possibilité a été prévue –et même cherchée– par le réalisateur qui, comme nous avons déjà montré, a produit toute une série de documents audiovisuels, oraux et textuels pour expliquer sa technique, et pour décrire comment il a fait ses films, particulièrement ceux faits avec la technique du son dessiné. Il s’agit donc d’une stratégie –dans le sens de Fontanille– dans laquelle *Synchromie* est une partie, peut-être la plus importante, pour exposer une découverte technique et une possibilité artistique.

Dans ce niveau de pertinence –la stratégie– un autre type d’énonciataire est construit : celui qui est conscient de la technique, dans le double sens de pratique et de dispositif : un énonciataire qui connaît –qui a vu– l’objet-support et la pratique du dessin sonore, et qui a audiovisionné les correspondances image/son dans les exemples didactiques profusément montrés dans les documentaires, les « modes d’emploi », les interviews, les « making of ». En bref, un mégaénonciataire qui a parcouru les vicissitudes de la création en suivant un magicien pas si silencieux qui révèle, peu à peu, toutes ses astuces.

## Chapitre 5. Les sons comme événements

*Movements achieve true freedom when they are released from the constraints of storytelling and acting. I believe there is something that can be only felt when there is a complete freedom of movement. This is the reason I persist in making non-narrative films.*

**Mirai Mizue**

Un des problèmes qui se pose avant d'aborder le type de films d'animation qu'on vient d'analyser est la question de leur « narrativité ». Dans les films de McLaren les personnages peuvent être des lignes, des formes ou des « barres » et leurs « actions » ne sont que purs mouvements ou transformations formelles ou chromatiques. En fait, ce sont souvent les sons qui leur donnent un certain ordre, un certain sens narratif. Le cinéma d'animation dit « expérimental » ou « abstrait » nous confronte avec un autre type de « contrat » entre sons et images et, peut-être, avec un degré particulier de narrativité.

Dans ce chapitre je voudrais développer cette question pour essayer de proposer quelques outils descriptifs pour aborder le son dans le cinéma d'animation, cette fois d'après la perspective des pratiques interprétatives du spectateur, et à partir de la notion de « dispositif animatique mental de réception ». Pour commencer, il est nécessaire de revoir les conditions qui peuvent être considérées comme « minimales » et suffisantes pour qu'il ait un récit. Il y a un ensemble de concepts récurrents qui apparaissent lorsqu'on veut définir ce « récit minimal » : action, transformation, événement, changement...

Néanmoins, dans les différentes approches de la narrativité, les sémioticiens et les narratologues ne sont pas toujours d'accord. Pour commencer par les derniers, dans le champ de la narratologie audiovisuelle il y a eu de nombreuses tentatives pour définir la narrativité à partir du concept d'événement. Par exemple André Gaudreault inclut l'événement dans sa liste des phénomènes qui sont à la base de la narrativité :

*...peut en effet être considéré comme narratif, on l'a dit plus haut, tout énoncé qui relate des actes, des gestes ou des événements ayant entre eux une « relation de succession » et développant un « rapport de transformation ». C'est sur la base de tels principes, fort peu exigeants au demeurant, que l'on peut être amené à conclure, comme je l'ai fait ailleurs, que*

*tout film, et même tout plan (tout plan mobile, à tout le moins), peut à juste titre être considéré comme un récit.* (GAUDREULT, 1999, pp. 44-45)

Gaudreault rappelle l'idée de Brémond selon laquelle, pour considérer qu'un message communique un récit il est nécessaire que « par ce message, un sujet quelconque (animé ou inanimé, il n'importe) soit placé dans un temps  $t$ , puis  $t + n$  et qu'il soit dit ce qui l'advient à l'instant  $t + n$  des prédicats qui le caractérisaient à l'instant  $t$  » (Bremond, 1973, pp. 99 - 100). Gaudreault corrige et complète cette définition :

*Un message est considéré comme narratif (...) lorsqu'il présente un sujet quelconque inscrit dans un quelconque procès de transformation, et qu'ainsi le sujet soit placé dans un temps  $t$ , puis  $t + n$  et qu'il soit dit ce qu'il advient à l'instant  $t + n$  des prédicats qui le caractérisaient l'instant  $t$ . On peut donc vraisemblablement avancer, sur la base de cette définition maintenant amendée, **que tout plan est un énoncé narratif (...)**, puisque par définition tout plan présente une suite (succession) d'images en mouvement (transformation).* (GAUDREULT, 1999, p. 49)

Gaudreault identifie alors « deux niveaux de récit au cinéma. Le premier de ces niveaux narratifs correspond au microrécit que chaque plan, en tant qu'énoncé photogrammatique, communique et qui constitue la base sur laquelle est engendré un second niveau narratif, supérieur, qui naît de la juxtaposition des microrécits communiqués par chaque plan » (GAUDREULT, 1999, p. 51).

Mais ces niveaux ne sont pas sans problèmes parce que, comme l'a bien souligné Roger Odin, toute succession d'images en mouvement (en transformation) est un microrécit, « le moindre mouvement : un œil qui s'ouvre, une feuille qui bouge, une goutte d'eau qui tombe est un microrécit et tout plan peut donner lieu à autant de microrécits que de mouvements repérables par le spectateur à l'écran, c'est-à-dire, sinon une infinité, du moins une multitude ». Et il s'en suit une autre idée indéfendable :

*Si un plan peut contenir une multitude de transformations, peut-on encore dénommer ces transformations des microrécits ? Quel spectateur accepterait de considérer le mouvement d'un œil qui s'ouvre ou celui d'une feuille qui bouge comme un récit même micro ?* (ODIN, 2000, p. 27)

Pour résoudre ces difficultés, Odin propose d'introduire la catégorie d'*action* comme ordonnatrice des mouvements et transformations, afin que ces mouvements et transformations

soient considérés comme constitutifs d'un niveau de narrativité. Nous ne pouvons pas oublier que pour Odin la narrativité est le résultat d'un mode de lecture, une opération qu'il appelle « narrativisation ». C'est dans ce sens qu'il affirme :

*Toute succession d'images en mouvement ne conduit donc pas à narrativiser ; si le passage du mouvement à la narrativité semble se faire de façon « naturelle », c'est parce que j'ai en moi le désir de narrativiser ; d'où l'illusion que c'est le film lui-même qui est narratif. Mais d'autres lectures restent possibles. (ODIN, 2000, p. 28)*

La narrativisation est, donc, une procédure interprétative parmi d'autres. En effet, Odin affirme que cette procédure ne s'active pas quand le spectateur lit un film comme abstrait : « Bien qu'il donne à voir de multiples transformations de formes et de couleurs, il n'y a pas d'actions dans un film lu comme abstrait (et s'il arrive qu'un film abstrait produise un sentiment de narrativité, c'est que je ne le lis pas réellement comme abstrait) » (ODIN, 2000, p. 28).

Par contre, un film est lu comme narratif quand le spectateur identifie ce que l'auteur appelle, en suivant Paul Ricœur, le *réseau conceptuel de l'action* : l'assignation au mouvement d'un but, des motifs, et des *agents* qui effectuent le mouvement et qui interagissent avec d'autres agents dans certaines circonstances. D'après Odin il y a une « seconde opération qui insère l'action dans un *système symbolique* qui lui « fournit un *contexte de description* » ; une action n'est lisible qu'une fois intégrée à un tel contexte ». Et une troisième opération « fait apparaître dans l'action 'des structures temporelles qui appellent la narration' (...) ; toute action peut être vue comme une sorte « d'histoire potentielle », c'est-à-dire, comme une sorte d'inducteur élémentaire de récit » (ODIN, 2000, p. 28).

Pourtant, toutes ces définitions tombent dans un problème déjà souligné par la sémiotique greimassienne par rapport à la conception de l'événement :

*(...) certains sémioticiens (et certains narratologues aussi, on peut ajouter), s'inspirant notamment des logiques de l'action, utilisent ce terme comme s'il désignait un donné simple et pour ainsi dire "naturel"; on voit, au contraire, que l'événement est une configuration discursive et non une unité narrative simple. (GREIMAS & COURTES, 1993)*

Dans la tradition greimassienne, pour que l'action d'un sujet soit considérée comme partie d'un événement, elle doit être « reconnue et interprétée par un sujet cognitif autre que le

sujet du faire lui-même et qui peut être soit l'actant observateur installé dans le discours (cf. le témoin), soit le narrateur, délégué de l'énonciateur » (Ibid.). D'où le fait que la dimension pragmatique soit appelée aussi « événementielle » dans la mesure où « s'y trouvent représentés et décrits les enchaînements de comportements somatiques » (ibid.). Et l'action est définie aussi dans le dictionnaire, « comme une organisation syntagmatique d'actes ». Et plus bas, on trouve : «...l'action est un programme narratif 'habillé', le sujet y étant représenté par un acteur et le faire converti en procès » (Ibid.). On peut continuer cette « navigation » dans le Dictionnaire en lisant l'entrée « Procès », pour trouver que cette précision « sert à désigner le résultat de la conversion de la fonction narrative de faire, conversion qui s'effectue grâce aux investissements complémentaires des catégories temporelles et surtout aspectuelles » (Ibid.).

Et cela nous renvoie à l'entité théorique de l'observateur, puisque « ...l'événement se définit comme l'action envisagée du point de vue de l'observateur. De même, les catégories aspectuelles ne s'expliquent que par la présence de l'observateur qui se prononce implicitement sur le faire du sujet au moment de sa conversion en procès » (Ibid.).

Ce qui manque donc dans les propositions des narratologues quand ils parlent de la narrativité à partir des notions d'action ou d'événement est le concept d'aspectualisation, formulé au début dans le dictionnaire :

*(...) l'aspectualisation d'un énoncé (phrase, séquence ou discours) correspond à un double débrayage : l'énonciateur déléguant dans le discours, d'une part un actant-sujet du faire, et de l'autre, un sujet cognitif qui observe et décompose ce faire en le transformant en procès (caractérisé alors par les sèmes de durativité ou de ponctualité, de perfectivité ou d'imperfectivité (accompli/inaccompli), d'inchoativité ou de terminativité). (Ibid.).*

Ainsi la question de l'aspectualisation est une partie fondamentale de la théorie de l'énonciation en général et de l'énonciation audiovisuelle en particulier. Néanmoins, la question de ces rapports exclusifs avec la narrativité n'est pas toujours claire, surtout si l'on se souvient des paroles de Roger Odin sur les lectures « abstraites » de certains films, lectures entendues comme « non narratives ». L'aspectualisation, donc, doit-elle être réduite aux procès de narrativisation, où fait-elle aussi partie des autres « modes » de lecture, ou même des autres « structures discursives » ? Et, de plus, le concept d'observateur est-il suffisant pour rendre compte de la complexité du discours audiovisuel ? Pour aborder ces questions je propose de ne pas abandonner le concept d'événement qui, à mon avis, couvre les

phénomènes normalement appelés « narratifs », mais aussi d'autres types de phénomènes – perceptifs, « naturels », « abstraits » –, qui ne sont pas lus comme narratifs, mais pourtant comme des *processus*. Mon idée est que la différence entre ces types de processus est due à différentes formes de l'aspectualisation des événements.

Mais avant de développer cette idée, je voudrais faire un petit détour par les théories cognitives de l'événement, afin de trouver des spécificités des discours audiovisuels en ce qui concerne l'aspectualisation.

### **La cognition d'événements et l'aspectualisation des événements audiovisuels**

Pour les théories cognitives, et notamment pour la psychologie cognitive, les événements jouent un rôle fondamental dans la compréhension du monde en général. Comme le disent Zacks et Tversky,

*Les événements sont ce qui nous arrive, ce que nous faisons, ce que nous attendons avec plaisir ou crainte, et ce que nous évoquons avec tendresse ou regret. Une grande partie de notre comportement est guidé par notre compréhension des événements. Nous percevons des événements lorsque nous observons le monde qui se déroule autour de nous, nous participons à des événements lorsque nous agissons sur le monde, nous simulons les événements sur lesquels on a entendu ou on a lu, nous utilisons notre connaissance des événements pour résoudre les problèmes. (RADVANSKY & ZACKS, 2014, p. 1)<sup>193</sup>*

Les exemples énumérés ici nous donnent une idée très générale des événements, y compris des phénomènes du monde naturel, mais aussi des événements « rapportés » dans les textes et les énoncés. Ces auteurs donnent aussi une définition plus précise, mais pas encore suffisante : un événement est « un segment de temps dans un emplacement donné qui est conçu par un observateur comme ayant un début et une fin »<sup>194</sup>. En ajoutant l'observateur, la théorie de la cognition des événements se rapproche de la théorie de l'aspectualité.

---

<sup>193</sup> “*Events are what happens to us, what we do, what we anticipate with pleasure or dread, and what we remember with fondness or regret. Much of our behavior is guided by our understanding of events. We perceive events when we observe the world unfolding around us, participate in events when we act on the world, simulate events that we hear or read about, use our knowledge of events to solve problems*”.

<sup>194</sup> “*...a segment of time at a given location that is conceived by an observer to have a beginning and an end*”.

Néanmoins, il manque encore des éléments qu'on trouve cependant dans les travaux de Barwise et Perry en sémantique situationnelle :

*Les composants basiques d'un événement sont des individus, les relations entre les individus, les propriétés des individus, les états de l'événement et des emplacements spatiotemporels. Ces composants sont présents dans tous les événements. Les individus peuvent être considérés soit comme des entités entières (par exemple, Joe), ou peuvent être décomposés en parties dont chaque partie est considérée individuellement comme un individu (par exemple, le bras de Joe). (RADVANSKY & ZACKS, 2014, p. 9)<sup>195</sup>*

Pour rendre compte de la cognition des événements, Radvansky et Zacks se basent sur la théorie des « modèles mentaux ». En l'appliquant à leur objet, ils parlent de « modèles d'événement », comme « des représentations que, croyons-nous, les gens utilisent pour penser les événements, en rassemblant l'information sur des gens, des objets, de l'espace, du temps, des causes, des buts, et le reste » (RADVANSKY & ZACKS, 2014, p. 16). Les auteurs proposent une hiérarchie des modèles mentaux avec une première distinction entre *modèles de systèmes* (system models), et *modèles d'événements* (event models). Les premiers représentent bien des systèmes fonctionnels comme, par exemple, le fonctionnement d'un dispositif mécanique (dans ce cas ils seront *modèles de systèmes physiques*) ou bien de systèmes « abstraits » (comme des théories mathématiques) ou qui ne peuvent pas être perçus directement, comme des théories de la physique subatomique. Dans ce cas, on parle de *modèles de systèmes abstraits*.

Les modèles d'événement, pour leur part, peuvent être des *modèles de situations* – aussi nommés *modèles discursifs*– ou *modèles d'expériences*. Les modèles de situation ou discursifs sont des modèles d'événements dérivés du langage, tandis que les modèles d'expérience sont des modèles dérivés « de l'expérience vivante, interactive » :

*Les modèles d'expériences sont des représentations d'événements provenant de l'expérience perceptive motrice, comme notre propre interaction avec le monde, la télévision, le cinéma et les technologies de réalité virtuelle. Certains chercheurs travaillant avec des récits ont pris*

---

<sup>195</sup> “The basic components of an event are individuals, the relations among individuals, individuals' properties, event states, and spatiotemporal locations. These components are present in all events. Individuals can either be considered as whole entities (e.g., Joe), or they can be broken down into parts in which each part is considered as an individual (e.g., Joe's arm)”.



*comme hypothèse de travail que les modèles de situation dérivés de la langue partagent la plupart de ces propriétés avec les modèles d'expériences. Nous pensons que ceci est une bonne hypothèse à prendre, mais il est important de la tester régulièrement. (RADVANSKY & ZACKS, 2014, p. 17)*<sup>196</sup>

Cette distinction est importante pour l'analyse du phénomène sémiotique qui m'intéresse ici, car l'introduction des phénomènes audiovisuels dans la catégorie d'événements d'expériences, permet d'établir quelques particularités des films qu'on va aborder plus bas (et des films en général). L'écart entre les textes verbaux et les expériences audiovisuelles peut être très petit ou très grand, selon la façon dont on le regarde. Il y a toujours des équivalences, des correspondances, mais elles doivent être toujours des hypothèses à tester<sup>197</sup>.

Comme pour Radvansky et Zacks, ce qui nous intéresse ici est donc la distinction entre les deux types de modèles d'événements. Et ce qui m'intéresse particulièrement dans cette approche est ce que les auteurs dénomment la *théorie de la segmentation des événements* (Event Segmentation Theory, EST)<sup>198</sup>. Par « *segmentation* » les auteurs comprennent le procès par lequel les gens fractionnent l'activité continue (ongoing activity) en événements discrets, et un modèle d'événement est construit pour chaque événement. La segmentation a à voir avec l'identification de *frontières de l'événement* (event boundaries), qui séparent un événement d'un autre.

D'après Radvansky et Zacks,

---

<sup>196</sup> “*Experience models are representations of events derived from perceptual-motor experience, such as our own interaction with the world, television, film, and virtual reality technologies. Some researchers working with narratives have taken as a working assumption that situation models derived from language share most of their properties with experience models*”.

<sup>197</sup> D'un point de vue cognitif, les différences ne sont pas radicales. Comme l'affirme Zacks : « ...les modèles d'événement construits à partir de films et à partir de livres sont plus semblables que différents. Pourquoi? Parce que l'architecture neuronale que nous possédons pour la construction de modèles n'a pas évolué pour nous aider à comprendre les films ou les romans: elle a évolué pour nous aider à traiter avec le monde réel » (...*event models from movies and from books are more alike than different. Why? Because the neural architecture we possess for model building did not evolve to help us understand movies or novels: It evolved to help us deal with the real world.*) (Zacks, 2015, p. 29).

<sup>198</sup> Je vais utiliser ici l'acronyme français équivalent : TSE.

*la TSE commence par le présupposé que les observateurs tentent de prédire l'avenir proche en tant que partie permanente de la perception. Le traitement perceptif transforme les entrées sensorielles en représentations élaborées qui comprennent des prévisions. Ce traitement est influencé par un modèle de travail de l'événement en cours. (RADVANSKY & ZACKS, 2014, p. 30)<sup>199</sup>*

Les prévisions sont plus difficiles quand il n'y a pas d'identification claire des limites ou frontières entre événements. Pour cette raison, le traitement cognitif fait particulièrement attention quand une limite est identifiée.

La segmentation des événements est dirigée en grande partie par des activités orientées par objectifs : « les buts offrent des contraintes majeures de causalité guidant les expectatives (...) Les buts offrent aussi les principales contraintes dans la structure de segmentation des actions humaines » (RADVANSKY & ZACKS, 2014, p. 47)<sup>200</sup>. D'une façon très proche de la description de la narrativité centrée dans l'action orientée par objectifs qu'on a trouvée dans la narratologie, la TSE affirme que les actions orientées par buts sont un des critères les plus forts pour la segmentation des événements, et qu'elles permettent de faire une segmentation hiérarchisée, et d'identifier, à partir d'objectifs et sous-objectifs, des événements enlacés et aussi hiérarchisés<sup>201</sup>. Grâce à cette hiérarchisation « les gens maintiennent simultanément des modèles d'événements sur plusieurs échelles de temps. Dans chaque échelle temporelle, il n'y

---

<sup>199</sup> “EST begins with the presupposition that observers attempt to predict the near future as an ongoing part of perception. Perceptual processing transforms sensory inputs to elaborated representations that include predictions. This processing is influenced by a working model of the current event”.

<sup>200</sup> “Goals provide major causal constraints on human behavior that guide expectations. (...) Goals also provide major constraints on the segmental structure of human actions”.

<sup>201</sup> De fait, dans un autre ouvrage Jeffrey Zacks propose que la compréhension narrative consiste principalement en la construction des modèles mentaux des événements du récit (Zacks, 2015, p. 26). À mon avis, cette théorie est compatible avec la théorie des pratiques sémiotiques et, de fait, peut être un complément intéressant pour les analyses des « cours d'action », dans la mesure où elle sert à rendre compte de l'interprétation des situations "en direct", soit des situations perçues dans le monde ou dans un texte ou un film. Dans ce dernier cas il s'agira de « traiter un texte ou une peinture sous le point de vue pratique, en focalisant le modus operandi de leur production et/ou de leur interprétation, et sans tenir pour pertinent le fait que ce texte ou cette peinture sont « achevés » » (Fontanille, L'analyse du cours d'action : des pratiques et des corps, 2011).

a qu'un seul modèle d'événement, mais à tout moment il y a des modèles d'événements actifs pour des échelles de temps multiples » (RADVANSKY & ZACKS, 2014, p. 52)<sup>202</sup>.

La TSE propose l'existence dans la mémoire à long terme, des *schémas d'événements*, c'est-à-dire, des *événements type* basés sur des expériences préalables et réitératives qui sont invoquées pour construire des modèles de travail d'une situation concrète dans un cours d'action. L'une des grandes questions soulevées par cette théorie est donc : la compréhension des événements est-elle toujours régie par des schémas ou des connaissances préalables sur cette situation (en termes cognitifs, un traitement *top-down*) ? Ou l'information sensorielle peut-elle aussi fournir des indices objectifs sur les limites d'un événement (traitement *bottom-up*) ?

Pour aborder cette question les chercheurs Bridgette M. Hard, Barbara Tversky, et David S. Lang ont développé des expériences sur l'interprétation des événements dits « abstraits », puisque les événements de ce type sont plus difficiles à rapporter avec des modèles préalables, de sorte que leur segmentation et hiérarchisation peut provenir directement des inputs perceptifs (*bottom-up*). Ces expériences ont été inspirées par Heider et Simmel qui en 1944 ont fait une étude dans laquelle

*des observateurs regardent un film d'animation sur une grande figure géométrique et deux petites qui se comportent selon un scénario de « chasse ». Beaucoup de gens ont interprété les mouvements des figures géométriques comme des actions intentionnelles anthropomorphiques. Ils ont décrit des histoires similaires dans lesquelles le grand triangle harcèle les petites figures. Ainsi, cette animation peut être interprétée selon un schéma familier d'événement hiérarchisé.* (HARD, TVERSKY, & LANG, 2006, p. 1223)<sup>203</sup>

---

<sup>202</sup> "... People simultaneously maintain event models on multiple timescales. At any given timescale there is only one event model, but at any time there are event models active for multiple timescales".

<sup>203</sup> "...observers watched an animated film of one large and two small geometric figures behaving according to a "chase" scenario. Many people interpreted the movements of the geometric figures as humanlike, intentional actions. They described similar stories involving the large triangle's bullying the two smaller shapes: chasing them and trying to capture them. Thus, this animation can be interpreted according to a familiar, hierarchically organized event schema".

Pourtant les événements dans cette animation sont abstraits, puisqu'il s'agit de figures géométriques sans visage, membres ou articulations, et la scène est elle aussi abstraite (il n'y a pas un environnement repérable).

Les auteurs ont fait une nouvelle expérience à partir de deux films animés, abstraits eux aussi. Le premier suit un scénario de « chasse », et le second un scénario de « cache-cache ». Les « protagonistes » des deux films sont des figures géométriques. Les chercheurs ont projeté les films de façons différentes : « vers l'avant pour certains observateurs et en sens inverse pour les autres (...) nous avons manipulé la familiarité des observateurs avec les événements à travers des visionnements répétés et l'instruction d'écrire un récit avant la segmentation » (HARD, TVERSKY, & LANG, 2006, p. 1223). Ces contraintes ont pour objectif d'offrir deux conditions de visionnement différentes : une condition de nouveauté et une condition de familiarité.

Les résultats de l'expérience sont très intéressants<sup>204</sup>, puisque la grande conclusion est que la structure des événements semble avoir une base physique, perceptive :

*Les données indiquent que les schémas d'événements fondés sur des objectifs jouent un rôle minime dans la détection de la structure de l'événement. Les participants ont sélectionné des événements similaires, organisés hiérarchiquement, indépendamment du fait qu'ils ont interprété ces événements comme des actions intentionnelles ou des mouvements physiques. Ces résultats suggèrent que la structure de l'événement a une base physique que les observateurs peuvent exploiter avant de construire des interprétations de l'événement. En accord avec cette explication, les participants qui ont segmenté les événements sur le premier visionnement ont sélectionné les mêmes limites des événements que les observateurs qui avaient vu les événements plusieurs fois et qui ont construit un récit pour eux. (HARD, TVERSKY, & LANG, 2006, p. 1229)<sup>205</sup>.*

---

<sup>204</sup> Pour plus de détails sur l'expérience, l'article complet peut être téléchargé à l'adresse: <http://link.springer.com/article/10.3758/BF03193267>

<sup>205</sup> *“The data indicate that goal-based event schemas play a minimal role in detecting event structure. The participants selected similar, hierarchically organized events regardless of whether they interpreted those events as intentional actions or physical movements. These findings suggest that event structure has a physical basis that observers can exploit before building event interpretations. Consistent with this explanation, the participants who segmented the events on the first viewing*

Cette base physique de la structure des événements est décrite d'après l'idée de Newton et al. selon laquelle les limites d'un événement pour l'action humaine coïncident avec des « éclats de changement » dans la position corporelle d'un agent: « Peut-être des observateurs d'événements abstraits utilisent-ils aussi des éclats de changement dans les positions des 'agents' géométriques pour détecter les limites de l'événement » (Ibid., p. 1229)<sup>206</sup>.

Les êtres humains peuvent donc identifier, segmenter et anticiper des événements même si les entités impliquées ne sont pas anthropomorphes ou ne sont pas associées avec des « actions » orientées par objectifs.

### **Objets, quasi objets et événements**

Dans une approche sémiotique en consonance avec les recherches cognitives, Douglas Niño intègre la théorie des événements en ce qu'il appelle une *ontologie agentive* : l'ensemble de choses, objets, procès, êtres, etc., auxquels les agents donnent un sens. Dans un certain sens, cette ontologie agentive est équivalente à la *sémiotique du monde naturel* de l'école de Paris<sup>207</sup>. La sémiotique agentive tire son nom des idées d' « agent », « agence » et « agenda ». Dans cette perspective, un agent est un être mobilisé dans le monde en conformité avec un objectif, son *agenda*. Un objet du monde (ou un procès, comme c'est le cas des événements) devient un « objet sémiotique » quand il est établi comme « la cible » d'un agenda dont le but est d'attribuer un sens à l'objet en question. Niño appelle ces éléments *items sémiotiques* (NIÑO, 2015, p. 170).

---

*selected the same event boundaries as did the observers who had seen the events multiple times and constructed a narrative for them”.*

<sup>206</sup> “Perhaps observers of abstract events also use bursts of change in the positions of the geometric “agents” to detect event boundaries”.

<sup>207</sup> Or une grande différence c'est que, pour être une sémiotique, la « grille » de lecture du monde naturel doit avoir un plan d'expression et un plan du contenu. Par contre, chez Niño l'ontologie agentive est distinguée de ce qu'il appelle la *sémantique agentive*, entendue comme « les conditions qui permettent d'attribuer un sens à l'ontologie, à son organisation, et l'effet attendu et effectif dans l'expérience » (Niño, 2015, p. 172). Cela ne veut pas dire que l'ontologie n'a à voir qu'avec la « réalité objective », puisqu'elle comprend toute entité qui puisse être conçue par les agents, y inclus les unicorns, les dragons et tous les objets produits par l'imagination humaine.

Dans cette ontologie agentive, les événements sont entendus à la manière de la TSE (comme un ensemble d'éléments ou entités en interaction, avec relations dynamiques) dans un espace et pendant un temps définis, et qui peuvent être segmentés en sous-événements ou « phases ». Mais, il y a des ensembles d'éléments situés et avec une durée, mais qui semblent « statiques ». Niño appelle *faits* ce type d'événements. Un événement peut être constitué donc par des (sous) événements ou par des *faits*. En d'autres termes, les éléments qui participent à un événement (ou à un fait) sont définis comme

- A. Assemblages unitaires de
  - 1. propriétés ou
  - 2. d'autres objets qui soutiennent entre eux des relations méronymiques, et qui sont
- B. reconnaissables et identifiables
  - 1. de façon multimodale ou intermodale, de mode direct ou
  - 2. de mode indirect (à travers des représentations);
- C. qui sont stables (par exemple en leur morphodynamique) ;
- D. qui peuvent être des guides pour des actes/actions, en raison de leurs fonctions (NIÑO, 2015, p. 178).

Le point E m'intéresse particulièrement, puisqu'il a à voir avec la sensorialité : qu'un objet soit reconnaissable de façon multimodale veut dire qu'une de ses propriétés peut être captée à travers différentes classes d'information provenant de systèmes perceptuels différents. Qu'un objet soit reconnaissable de façon intermodale veut dire qu'une de ses propriétés peut être captée en obtenant *la même* information à travers des systèmes perceptuels différents. Par exemple, la reconnaissance de « la fraîcheur du pain » peut être obtenue de façon multimodale à travers son craquement (information auditive), sa température ou sa résistance tactile ou même par son apparence quand on le coupe. En outre, l'identification de la forme d'un objet peut être réussie à travers la vue ou le toucher. Il s'agit donc d'une reconnaissance *intermodale* d'un même type d'information : la forme (NIÑO, 2015, p. 182).

D'après Niño, certaines propriétés des objets peuvent être individualisées et assumées comme des entités indépendantes (par exemple, une mélodie). Cet auteur appelle *quasi-objets* ce type de propriétés (monomodales) individualisées. Quand l'item sémiotique qui capture notre attention est un événement (constitué par un ensemble de sous événements ou phases),

on est face à une *narration*. Et quand l'événement en question est constitué par des propriétés individualisés comme quasi-objets, on est donc face à une *quasi-narration* (NIÑO, 2015, p. 401). L'auteur donne l'exemple des films abstraits de Norman McLaren, « qui ne représentent pas les « événements », mais qui semblent des transformations dynamiques de propriétés ou quasi-objets » (NIÑO, 2015, p. 401).

On peut se demander : qu'est ce qui se passe quand on réifie une propriété non comme un objet, mais comme un événement ? Cela nous permet d'introduire d'autres phénomènes perceptifs dans la catégorie, maintenant élargie, de l'événement : les sons.

### **Les sons comme événements**

Dans le chapitre 1, on a déjà mentionné que, d'un point de vue phénoménologique (partagé en grande mesure par la sémiotique tensive), les sons sont perçus comme événements (IDHE, 2007, p. 53). Cela veut dire que les limites d'un son (ce qui le donne sa « forme ») sont essentiellement temporelles. D'après cette perspective, le monde sonore est un « flux » continu et temporel, et les événements sonores sont « coupés » par l'écoute en identifiant leur temporalité. Comme l'a dit aussi Casey O'Callaghan, « intuitivement, les sons sont des événements (happenings) qui ont lieu dans notre environnement. Cela est évident dans la langue que nous utilisons pour parler de sons. Les sons, comme des explosions et des concerts, se produisent, ont lieu, et ainsi de suite. Couleurs, formes et violons ne le font pas. » (O'CALLAGHAN, 2007, p. 57)<sup>208</sup>.

Ce philosophe souligne aussi des différences entre les limites des objets et les limites des sons : « Contrairement aux objets, dont les limites spatiales sont relativement claires, mais les limites temporelles sont plus difficiles à discerner, les événements sont caractérisés par des limites temporelles claires, mais des limites spatiales qui peuvent être obscures » (Ibid., p. 58). En outre, O'Callaghan affirme que les sons des *relata* causales, dans la mesure où on les associe avec leurs « sources » ou « causes ».

Cela lui permet proposer que les sons sont des « perturbations » (disturbings) dans le sens qu'ils sont « des événements dans lesquels un objet en mouvement perturbe un milieu

---

<sup>208</sup> “*Sounds, intuitively, are happenings that take place in one's environment. This is evident in the language we use to speak of sounds. Sounds, like explosions and concerts, occur, take place, and last. Colors, shapes, and fiddles do not*”.



environnant et le met en mouvement. Les heurts et percussions ne sont pas les sons, mais sont les causes de sons »<sup>209</sup> (Ibid., p. 61). Ces perturbations sont causées par « les activités des objets ou corps en interaction » (p. 64). La caractérisation des sons comme événements est due surtout à leur dimension temporelle (sa durée), ces emplacements distaux stables, le milieu qu'ils « perturbent » et ces rôles dans des lignes de causalité. Ces traits permettent d'individualiser les sons.

À ce point il est nécessaire de remarquer que les sons sont des événements au moins en deux sens : ils sont des événements physiques complexes, qui se déroulent indépendamment d'être perçus (et avec des dimensions que ne peuvent pas être perçues). Mais ils sont aussi des événements *perceptifs*, avec une configuration déterminée par les organes, le traitement et l'interprétation des êtres percepteurs. Comme l'a souligné Jean-François Bordron :

*...les sons ainsi nommés paraissent indépendants de la perception qui les accueille, mais aussi, dans une large mesure, les configure. Comme l'a enseigné Pierre Schaeffer, il existe des types bien différents d'écoute de telle sorte qu'il semble peu pertinent de décrire un son par sa seule analyse physique. La perception par ailleurs n'est pas statique. C'est une action, donc une dynamique dont le profil n'est pas nécessairement linéaire.* (BORDRON, 2010).

Don Ihde nous a fourni déjà le modèle husserlien pour la compréhension de la temporalité en l'appliquant aux sons : ce qui nous permet d'identifier la structure temporelle d'un événement (sonore) est la relation entre le point (focus) temporel de l'attention dans l'écoute, et les retentions et les protentions qui nous permettent de faire des anticipations du cours de l'événement (IDHE, 2007, p. 89)<sup>210</sup>.

Compte tenu de ce que la TSE et la théorie des événements en général cherchent à établir des limites, les bornes, et les « contours » des événements comme partie fondamentale de la cognition, mais aussi la « granularité » du traitement des événements (ses degrés,

---

<sup>209</sup> “[Sounds] are events in which a moving object disturbs a surrounding medium and sets it moving. The strikings and crashings are not the sounds, but are the causes of sounds.

<sup>210</sup> Dans sa proposition de sémiotique de la musique, Eero Tarasti fait une adaptation du modèle husserlien de la temporalité, pour rendre compte de la « rétention d'événements musicaux » (TARASTI, 1994, p. 64).

chevauchements, etc.), on peut les prendre comme un cadre théorique qui soutient l'idée selon laquelle l'aspectualité est à la base de la narrativité.

### **Retour à la sémiotique : l'aspectualité des événements sonores**

Comme on l'a dit au début, la théorie de l'aspectualité donne un cadre explicatif à la théorie des événements comme fondement de la narrativité. Néanmoins, à partir du « détour » cognitif, on a trouvé que l'identification, la segmentation et le traitement des événements peuvent se dérouler sans l'intervention des schémas complexes d'actions ou de comportements typifiés. Cette découverte est proche de l'idée de Fontanille et Zilberberg selon laquelle il est nécessaire d'inverser la relation structurale entre narrativité et aspectualité, pour prendre la dernière comme le terme qui détermine la première (FONTANILLE & ZILBERBERG, 2004, p. 347). Une idée similaire se trouve aussi chez Fontanille, quand il affirme :

*L'aspectualisation fait appel à un principe continu et tensif qui n'appartient pas en propre au niveau discursif, mais qui doit être postulé dans la théorie sémiotique dès les préconditions de la signification, et antérieurement à la catégorisation elle-même.* (FONTANILLE, 1991, p. 142).

L'installation de la théorie de l'aspectualité au niveau tensif, me permet d'aller un peu plus loin et d'affirmer que l'aspectualisation donne lieu à la conception des autres modalités discursives au-delà de la narrativité, ou au moins à différents degrés de narrativité dont les premiers peuvent être « lus » ou interprétés comme abstraits.

Je voudrais développer cette idée à partir de l'analyse d'un ensemble de courts-métrages animés qui ont comme sujet une des plus fortes propriétés des sons : le timbre.

### **Mirai Mizue, animation abstraite et narrativité**

Comme on a vu dans le chapitre précédent, l'animation abstraite a été une tendance très commune dans le panorama du milieu du XX<sup>e</sup> siècle, mais plus difficile à trouver au XXI<sup>e</sup> siècle. Pourtant il y a encore des animateurs qui ont continué la recherche sur les rapports son/image en utilisant l'animation abstraite comme moyen d'expression. Un des plus intéressants est l'artiste japonais Mirai Mizue.

Né à Tokyo en 1981, Mizue a fait ses études d'animation au département de design graphique de la Tama Art University. Son travail se caractérise par le dessin à la main, avec des stylos sur papier (parmi d'autres techniques), en faisant la mise en mouvement sans l'intervention de l'ordinateur. Néanmoins, il s'appuie sur la technologie numérique pour compléter et diversifier les compositions, y compris la bande-son. Dès son premier film, *Fantastic Cell* (2003), il est reconnu pour ses figures et personnages qui semblent faits de petites cellules graphiques (voir illustration 1), un motif visuel qui se répète dans la plupart de ses films, et sur lequel il a dit :

*Quand je fais un film d'animation, mon intention première n'est pas de raconter une histoire, mais de procurer au spectateur une certaine expérience à travers des mouvements et du métamorphisme. En vertu du processus de la division cellulaire, la cellule se prête très bien à cet objectif. L'échelle de la cellule est microscopique, donc elle donne une impression de voyager vers les espaces intérieurs du corps.* (Mizue, Interview, 2010)<sup>211</sup>.

Dès ce premier film, donc, la question de la narrativité est donc formulée par l'artiste. En effet, en prenant parti pour l'abstraction, Mizue s'inscrit dans une des polarités que j'ai utilisée aussi pour classifier l'animation, mais sans la problématiser. C'est maintenant le moment de le faire.

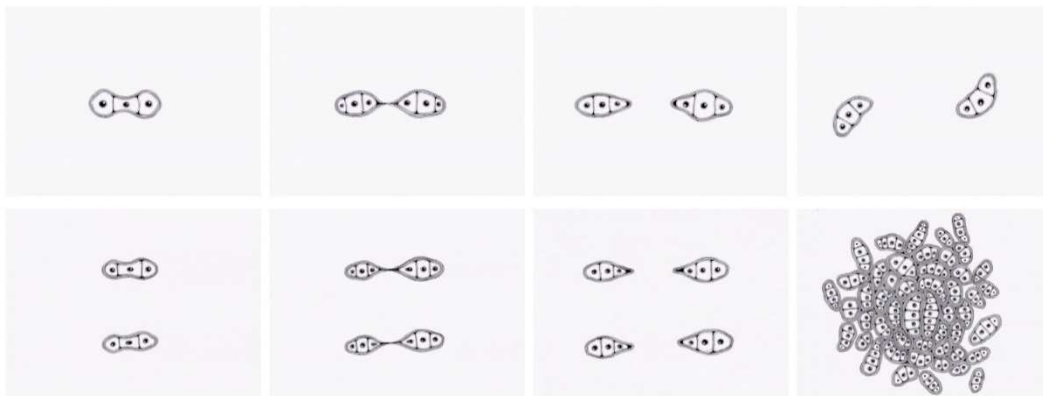


Figure 33 : Photogrammes du film *Fantastic Cell*

Source : [Mizue, 2003]

<sup>211</sup> « *When I make an animation film, my primary intention is not to tell a story, but to make the viewer experience something through movements and metamorphism. By virtue of the process of cell division, the cell lends itself very well to this objective. The scale of the cell is very microscopic, and so it imparts a feeling of traveling toward the inner spaces of the body* ».

Ainsi que l'a souligné Karen Beckman, dans son introduction au volume *Animating Film Theory* :

*Tant la théorie comme l'histoire du cinéma souvent se basent sur une série de termes binaires, y compris continu versus non continu, narratif versus expérimental, indicielle versus à la main, et animé versus pris en vue réelle. Bien que ces oppositions puissent être utiles, elles peuvent aussi conduire à des présomptions inexactes dans la mesure où elles ne reflètent pas toujours avec précision les différences contenues dans l'un ou l'autre de ces termes, comme expérimentale ou narrative. (BECKMAN, 2014, pp. 1-2)<sup>212</sup>.*

Cette auteure, qui veut montrer les difficultés auxquelles conduit le fait d'exclure l'animation de la théorie du cinéma, formule aussi une distinction entre le *changement* et le *mouvement* (associé avec l'action) pour remarquer que le second terme n'est pas le seul trait qui définit l'animation<sup>213</sup>. Cette distinction est pertinente dans la mesure où il y a des films abstraits qui ne présentent que des transformations ou changements que l'on peut difficilement définir comme des mouvements : un changement de couleur, par exemple. Il y a aussi des mouvements qui ne font pas partie d'une action ou ne sont pas le résultat d'une action. Finalement, on peut aussi se demander si *changement* et *transformation* sont la même chose, puisque le deuxième terme semble être plus orienté... (se transformer *en* une autre chose, ou avec tel ou tel objectif...). Ainsi dans l'animation dite abstraite on va trouver tout type de phénomènes qui oscillent entre toutes ces distinctions. Et le panorama devient plus compliqué si l'on considère les phénomènes sonores et leurs rapports avec l'image.

---

<sup>212</sup> « *Film theory and history both frequently rely on a series of binary terms, including continuous versus non-continuous, narrative versus experimental, indexical versus handmade, and animated versus live action. Though these oppositions can be useful, they can lead to inaccurate presumptions in that they do not always accurately reflect the differences contained within any one of these terms, such as experimental or narrative* ».

<sup>213</sup> Ou même le cinéma en général. Beckman cite aussi Peter Kubelka, qui affirme : « *le cinéma n'est pas mouvement. Le cinéma est une projection d'images statiques –c'est-à-dire, des images qui ne bougent pas- à un rythme très rapide. Et vous pouvez donner l'illusion du mouvement, bien sûr, mais c'est un cas spécial* » (Beckman, 2014, p. 3). [Cinema is not movement. Cinema is a projection of stills –which means images which do not move- in a very quick rhythm. And you can give the illusion of movement, of course, but this is a special case].

En retournant à Mirai Mizue, il est évident qu'il n'est pas ignorant de ces questions. Depuis le début de ses recherches esthétiques, l'animateur japonais a montré un grand intérêt pour trouver des correspondances et interactions entre ses dessins animés et la bande-son. En effet, dans *Fantastic Cell* il a utilisé la musique du *Valse des Fleurs*, du *Casse-noisette*, inspiré par des films de Disney comme *Le Vieux Moulin* (1937) ou *Fantasia* (1940)<sup>214</sup>.

Ce premier film est intéressant du point de vue des pratiques sémiotiques : réalisé lorsque Mizue était encore étudiant, l'animateur n'a pas obtenu les droits pour l'usage commercial de l'œuvre de Tchaïkovsky. Sur le DVD qui recueille ses films de 2003 à 2010, le court-métrage apparaît sans bande-son. Néanmoins, le musicien Pierre Slinckx a composé une bande sonore alternative pour le film, publiée aussi dans son canal Vimeo le 27 septembre 2013. Il s'agit d'une composition de musique électronique qui suit strictement (autant que possible) le taux d'apparition des images de Mizue. Le résultat est une combinaison audiovisuelle de style très contemporain (les créatures de Mizue semblent résonner frénétiquement avec les bruits électroniques de Slinckx), mais complètement différent de la fluidité que l'animateur a réussi à produire, en suivant la mélodie de Tchaïkovsky.

Dans ce dernier cas, c'est l'animateur qui a réalisé sa composition visuelle en s'adaptant à une composition musicale préalable. Dans le cas de Slinckx, c'est le musicien qui s'est adapté aux images de l'artiste japonais. On peut dire que les deux versions montrent une synchronisation plus ou moins précise, mais l'effet global est bien différent. Les différences mériteraient une analyse plus détaillée, mais ce qui nous intéresse maintenant dans ce film et ces deux versions est une caractéristique commune à l'ensemble du travail de Mirai Mizue : ses œuvres ne sont pas des produits fermés ; il y a toujours la possibilité de trouver un

---

<sup>214</sup> Quand quelqu'un a comparé *Fantastic Cell* avec l'œuvre d'Oskar Fischinger, Mizue a dit : « Je ne connaissais pas son travail au cours de la réalisation de ce film. Au contraire, je connaissais les films de Disney, tels que *Fantasia*, *Le Vieux Moulin* ou *Les Instruments de musique* (que j'ai vu quand j'étais très jeune), en raison de la synchronisation de la musique avec l'animation qui est typique de ces films » (*I wasn't familiar with his work during the making of this film. Rather, I was aware of Disney's films, such as Fantasia, The Old Mill or Toot, Whistle, Plunk, and Boom (Which I saw when I was very young), because of the synchronization of music and animation that were typical of these films*) (Mizue, Note sur *Fantastic Cell*, 2010). Fait intéressant, Disney a engagé Oskar Fischinger pour concevoir la séquence de la *Toccata et fugue en ré mineur* qui fait partie de son film *Fantasia*. C'est-à-dire que Mizue connaissait indirectement le travail de Fischinger... Sur Fischinger et sa participation à *Fantasia* voir (MORITZ, 2004, pp. 89-108).

personnage, une séquence ou même un film entier qui pourraient faire partie d'un nouveau projet.

Dans ses films postérieurs, Mizue a travaillé avec les compositions de différents musiciens et concepteurs sonores, parmi lesquels : Alice Nakamura (*Trip!-Trap!*, 2005 ; *Lost Utopia*, 2007), Kai&Co (*Jam*, 2009), Twoth (*Tatamp*, 2011), Toru Matsumoto (*And And*, 2011) Pascals (*Wonder*, 2012), ou Shugo Tokumaru (*The Baby Birds Of Norman McLaren*, 2014).

En général, le composant visuel des films de Mizue peut être conçu à la lumière de la sémiotique plastique, soit dans la version « école de Paris » (FLOCH, 1985), (GREIMAS A. J., *Sémiotique figurative et sémiotique plastique*, 1984) ou dans la version « école de Liège » (GROUPE.μ, 1992). En effet, quand il a été interrogé sur ses influences artistiques, l'animateur japonais a mentionné le peintre catalan Joan Miró: « Je peux sentir le rythme et la musique dans ses peintures. Bien que ce ne soient que des peintures, vous pouvez discerner le passage du temps en elles. Je veux créer le même rythme dans mon animation. Je tire une grande partie de mon inspiration de la peinture abstraite »<sup>215</sup>.

Pourtant, les théories sémiotiques du plastique manquent de concepts et d'instruments conceptuels pour rendre compte de la composante temporelle et dynamique des images, que ce soient des images statiques ou des images en mouvement. Dans un travail récent, César Díaz a essayé de compléter la perspective du Groupe μ en adaptant les trois sous-systèmes du signe plastique (couleur, forme et texture) aux régimes technologiques de l'image mouvement, et en ajoutant un quatrième système pour rendre compte de la temporalité et du mouvement. Ainsi, la texture est associée au « régime de la résolution de l'image » ; la couleur au « régime de la réponse de la lumière » et la forme au « régime du format ». Ce sont les aspects « tabulaires » de la photographie au cinéma. Les aspects temporels et dynamiques de l'image mouvement sont compris sous le « régime du taux de photogrammes » (DIAZ, 2015, pp. 96-134).

Cette proposition cherche à rendre compte des aspects plastiques de la photographie au cinéma et elle est de grande utilité pour décrire des phénomènes comme les ralentissements,

---

<sup>215</sup> « *I can feel the rhythm and music in his paintings. Although they are just paintings, you can discern a passage of time in them. I want to create the same rhythm in my animation. I draw much of my inspiration from abstract painting* ».

les accélérations et, en général, les aspects de la « vitesse » des images au cinéma, mais aussi les phénomènes relatifs aux orientations des images. Néanmoins, le « régime du taux de photogrammes » est une question relative à la « capture » mécanique du mouvement dans le dispositif photographique. Il faudra voir les possibilités qui s'ouvrent dans l'animation dans la mesure où le contrôle des photogrammes est ouvert à d'autres manipulations, dans ce que nous avons appelé le dispositif *animatique*, où la manipulation du mouvement et du temps a d'autres possibilités.

À cet égard, Patrick Barrès apporte des éléments importants à la compréhension du composant dynamique des images plastiques dans l'animation. Ainsi, par exemple, il nous rappelle l'idée deleuzienne des « espaces quelconques » et leur prolifération dans le cinéma expérimental, pour montrer comment « de la peinture au cinéma, nous passons par exemple de la « couleur-surface » à la « couleur-mouvement » » (Barrès, 2006, p. 24). En suivant cette idée, cet auteur affirme :

*Le cinéma d'animation centré sur le chantier de création plastique imprime dans le mouvement du film le libre-jeu des opérations plastiques, les traces gestuelles et les différents états de la matière, les accidents et les repentirs. Vu sous cet angle, il constitue un cinéma de recherches, d'expérimentations sur des paramètres plastiques (le support, le cadre, la surface, le décor, le mouvement, etc.) et sur les relations entre les paramètres (suivant des opérations de disjonction, de fragmentation, d'impression, d'accélération ou de ralenti, de multiplication, etc.) (BARRÈS, 2006, pp. 25-26).*

Barrès étudie avec rigueur les transformations et possibilités d'expérimentation dans le cinéma d'animation, dans le croisement des espaces de la création plastique (qui connecte l'animation avec les arts plastiques en général) de l'architecture et de la mise en scène théâtrale, et de l'image cinématographique, en les intégrant sous deux catégories : la *géométrie mobile* et l'*image-mouvement*. Il propose deux tableaux qui résument les critères classificatoires de ces catégories qu'il décrit de la façon suivante :

*Le premier tableau correspond à un investissement plastique local, en référence à un système global prédéfini et normé (le plan, le cadre, le centre géométrique). Ce régime de contraintes, caractéristique de l'espace du tableau, impulse une plastique du montage dans laquelle se repèrent des alternances, des continuités ou des ruptures. Nous pourrions, à la limite, transférer quelques-unes des opérations repérées par Christian Metz dans les catégories de sa « grande syntagmatique » à cet univers plastique.*



(...)

*Le second tableau correspond à un investissement plastique local instaurateur d'une globalité, une globalité construite, caractérisée par l'expression de nouvelles logiques plastiques : fragmentation (considérée comme travail en tension de la rupture et de la liaison), recouvrement (suivant le double jeu de recouvrir et de recouvrer). Ce deuxième système rejoint notamment les propositions des artistes du début du XX<sup>e</sup> siècle. (BARRES, 2006, p. 120).*

Tableau 4 : A et B chez Barrès (BARRES, 2006, pp. 118-119)

Tableau A

<b>GEOMETRIE MOBILE</b>		<b>Image par image</b> <i>Données spatiales</i>	<b>Flux d'images</b> (synthèse) <i>Données temporelles</i>
<b>Caractéristiques plastiques</b>	Éléments plastiques	Lignes, formes géométriques de base	Géométrie mobile
	Opérations plastiques	Répétition, multiplication, combinaison, élision, extension, érosion	Combinaison, transformation, répétition, multiplication, élision, extension, érosion
	Organisations plastiques locales	Motifs formels	Rythmes formels
<b>Systèmes de structuration ou de régulation globaux</b>		Système de coordonnées du plan et du cadre, exercice de centration ou de décentration	Alternances, contrastes, système de variations, convergence/divergence* clignotements

Tableau B

<b>IMAGE-MOUVEMENT</b>		<b>Image par image</b> <i>Données spatiales</i>	<b>Flux d'images</b> (synthèse) <i>Données temporelles</i>
<b>Caractéristiques plastiques</b>	Éléments plastiques	Lignes, traits, taches et plages de couleur, formes anamorphosées	Matériaux couleur et texture graphique en action
	Opérations plastiques	Impression, superposition, grattage, empâtement	Transformation, répétition, érosion, dilatation, extension, concentration, apparition/disparition, accélération <sup>206</sup>
	Organisations plastiques locales	Motifs chromatiques texturaux	Rythmes plastiques liés aux matériaux
<b>Systèmes de structuration ou de régulation plastique globaux</b>		Homogénéité/hétérogénéité de la surface, pouvoir absorbant, recouvrement, fragmentation	Gradations <sup>207</sup> , caractère absorbant, processus de variations, expansion/contraction, Attraction/répulsion <sup>208</sup>

Cette proposition est centrée sur la composante visuelle de l'animation, bien qu'elle prenne en compte des aspects rapportés aux rythmes formels et plastiques ; il s'agit toujours

des rythmes visuels –et même si on « descend » au niveau figural de la perception<sup>216</sup> (comme le fait Mizue lui-même dans son interprétation dynamique de l'œuvre de Jean Miró) on peut utiliser les mêmes catégories tensives pour parler des rythmes visuels et sonores–, il manque encore la pris en compte de la composante sonore dans la description des expériences plastiques<sup>217</sup>.

En tout cas, je partage l'idée de Barrès selon laquelle est possible de postuler « un processus narratif d'ordre plastique » où même d'autres types de processus plastiques dynamiques. L'œuvre de Mirai Mizue me permettra d'avancer un peu dans ce sens. Selon les propres mots de l'animateur japonais :

*Les mouvements atteignent la vraie liberté quand ils sont libérés des contraintes de la narration et de la comédie. Je crois qu'il y a quelque chose qui peut être ressenti seulement quand il y a une totale liberté de mouvement. Ceci est la raison pour laquelle je persiste à faire des films non narratifs. (Mizue, Interview, 2010)<sup>218</sup>.*

Néanmoins, cette « totale liberté de mouvement » est en réalité, dans presque tous ses films, soumise aux contraintes de la composante sonore, de sorte qu'il y a aussi chez Mizue une expérimentation plastique avec cette composante. On va voir dans la série *Timbre* qu'on peut parler aussi d'une *conception plastique du son*, et aussi d'un développement plastique de l'image à partir du son.

### **Les timbres de Mizue : « créatures » et événements**

La série de films qui nous intéresse dans ce chapitre n'est pas un de ses travaux collaboratifs, mais un de ses projets en solitaire<sup>219</sup>, au moins au début. Il s'agit de *Timbre A-Z* (2011), une expérience créative qui consiste à publier un film par jour pendant un certain

---

<sup>216</sup> À cet égard, voir par exemple l'ouvrage de Veronica Estay Stange, où elle propose –bien que dans un sens diachronique– la possibilité de placer à la base de la création artistique (dans les beaux-arts et la littérature) la *musicalité* (Estay Stange, 2014).

<sup>217</sup> Il le mentionne, mais il ne le développe pas voir (Barrès, 2006, pp. 19, 26 et d'autres).

<sup>218</sup> “*Movements achieve true freedom when they are released from the constraints of storytelling and acting. I believe there is something that can be only felt when there is a complete freedom of movement. This is the reason I persist in making non-narrative films*”.

<sup>219</sup> En fait, Mizue a assuré aussi la conception sonore de plusieurs de ses films.

temps<sup>220</sup>. Dans ce cas, le critère pour le nombre de jours a été l'alphabet, comme l'indique le titre du film ; c'est-à-dire, 26 films publiés dans le canal Vimeo de l'artiste<sup>221</sup> : le premier film de la série, *Timbre A*, a été publié le 19 janvier 2011 et le dernier, *Timbre Z*, le 13 février. Cette série a été intégrée dans un seul film : *Timbre A to Z*<sup>222</sup>, mais plusieurs des éléments de ces films ont été réutilisés dans autres productions, notamment le film *A long day of timbre* (publié le 22 février 2011), ou *Tatamp* (2011), codirigé par Shinichi Suda et, comme on l'a déjà indiqué, avec musique additionnelle de Twoth.

Dans la série originale, chaque film a une durée de 30 secondes, avec une complexité visuelle et sonore très diverse. Le sujet de la série est le « timbre », donc chaque film présente un timbre particulier (ou un ensemble de timbres) en couple avec une représentation visuelle dynamique.

Mais le *timbre* n'est pas un sujet facile, et Mizue le sait. Dans la théorie musicale, par exemple et comme l'a souligné Pierre Schaeffer depuis longtemps, la description du timbre mélange souvent les aspects physiques, perceptifs et psychologiques en produisant plus de confusion que de clarté (SCHAEFFER, 2003 [1966], pp. 93-94).

Selon la définition la plus générale, le timbre est la « propriété » qui caractérise un instrument, ou dans un sens plus large, un son. Mais les sons que nous entendons sont complexes, à savoir qu'ils sont composés de plusieurs ondes simultanées, puisque le mouvement vibratoire qui produit le son est toujours composé par un mouvement principal et des mouvements secondaires. Et bien que nous les percevions comme un seul phénomène, le timbre dépend de la quantité d'harmoniques dans un son et de l'intensité respective de chaque harmonique. Dans ce sens, le timbre n'est pas une propriété des sons, comme la tonalité (fréquence), l'intensité (amplitude) ou la durée, mais une combinaison de ces propriétés.

Entendu comme l'« identité » d'un son, le timbre peut être décrit aussi comme un événement, et, donc, être aspectualisé. En effet, les différentes approches théoriques du timbre offrent aussi différentes aspectualisations des timbres. Dans la synthèse analogique des sons,

---

<sup>220</sup> Ce n'est pas la seule fois que Mizue a fait ce type de projet. On peut mentionner la série *Kubrick*, dans laquelle l'artiste a publié un film par jour du 1er février au 14 février 2012 ; ou *Wonder*, une série de 365 jours/films à partir du 1er avril, 2012 jusqu'au 31 mars 2013.

<sup>221</sup> <https://vimeo.com/user5743956>

<sup>222</sup> Ce film est disponible à l'adresse suivant : <https://vimeo.com/31332018>

par exemple, les timbres sont décrits à partir de l'amplitude temporelle de l'énergie d'onde, c'est-à-dire son « enveloppe sonore », ce qui donne l'aspectualisation suivante :

Attaque (*attack*) → Chute (*decay*) → Maintien (*sustain*) → Extinction (*release*)<sup>223</sup>

Chez Schaeffer il y a aussi une théorie de l'aspectualisation du timbre dans sa description de la « forme temporelle » dans le chapitre VII de son *Traité*. Ce musicologue a été intéressé par l'attaque (l'inchoativité du timbre, en termes sémiotiques), et il a fait des expériences très intéressantes à ce sujet : à partir des enregistrements magnétiques d'une note grave de piano, Schaeffer et son équipe ont coupé les premières millisecondes (correspondant, à son avis, à l'*attaque*) et ont joué la bande à nouveau après la coupe. Ils ont été alors surpris par le fait que la note écoutée avait la même attaque, comme si elle n'avait pas été coupée. Alors Schaeffer a étendu ces expériences à différents types d'instruments. Il a conclu que « la perception musicale de l'attaque était en corrélation, d'une part, avec la dynamique d'ensemble du son, à savoir l'évolution énergétique, et d'autre part avec le contenu harmonique » (SCHAEFFER, 2003 [1966], p. 133).

Selon cette découverte, Schaeffer affirme que tous les sons ont généralement trois phases temporelles :

- Une phase d'établissement A;
- Une phase de maintien B;
- Une phase d'extinction C.

Ces phases ne sont pas complètement séparées. D'après Schaeffer, la dynamique générale du timbre est marquée par la vitesse d'établissement du son (phase A)<sup>224</sup>. Il y aura donc trois ordres de magnitude :

---

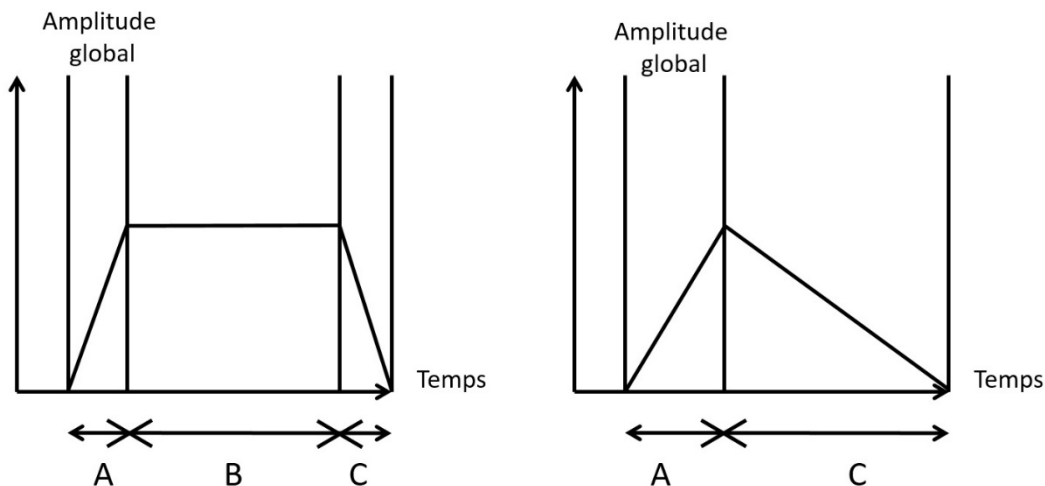
<sup>223</sup> Ce sont « les quatre paramètres d'un générateur d'enveloppe, partie essentielle d'un synthétiseur analogique. L'*attack* correspond à la durée d'attaque de la note, et le *decay* au temps qu'elle met pour retomber de son niveau maximal dans le domaine de *sustain*. Le *sustain* fixe le niveau de *sustain*, auquel la note se maintient tant que la touche reste enfoncée. Le *release* indique la durée d'extinction de la note une fois la touche relâchée » (DE NANTEUIL (ed.), 2008, p. 227). En effet, la majorité des *Timbres* de Mirai Mizue sont des sons *synthétiques*, produits par des instruments électroniques.

<sup>224</sup> La vitesse de l'attaque est complétée par ce que Schaeffer dénomme la *couleur* de l'attaque qui peut être « riche » ou « pauvre ».

- Établissements très rapides.
- Établissements de moyenne durée.
- Établissements longs.

D'où il émerge une série de lois de la perception des attaques:

1. Dans les sons soutenus, l'oreille est sensible en général pour qualifier leur perception de la vitesse de l'attaque, selon la façon dont l'énergie sonore apparaît dans le temps.
2. Dans les sons d'attaque percutée ou pointillée, suivie de résonance, l'oreille est plus sensible pour évaluer la perception de la vitesse de l'attaque quand l'énergie disparaît plutôt que quand elle commence (SCHAEFFER, 2003 [1966], pp. 135-138)



Figures 34 et 35 : Phases dynamiques des sons maintenus et Phases dynamiques de la percussion-résonance

Source : [SCHAEFFER, 2003 [1966], p. 134]

Dans cette perspective, l'inchoativité est le critère qui détermine tout le processus de la perception sonore dans l'aspectualisation du timbre. Pourtant, la durée du son détermine l'importance de l'attaque. Ainsi, chez Schaeffer c'est dans les sons brefs, comme les percussions, que l'attaque joue un rôle décisif pour caractériser les timbres ; pour les sons de moyenne durée, l'importance de l'attaque diminue et l'attention se centre plus dans l'évolution du son ; pour les sons soutenus avec *vibrato* (le cas le plus fréquent) le rôle de l'attaque est presque négligeable (SCHAEFFER, 2003 [1966], p. 132). Dans l'illustration 1, on peut regarder la structure aspectuelle complète des sons maintenus, tandis que l'illustration

2 montre la structure aspectuelle réduite des sons percussifs, où la durativité est réduite à un point<sup>225</sup>.

Chez Mirai Mizue on trouve tout un transcodage de ces aspects du timbre à travers des images animées. Nous allons maintenant présenter une description du fonctionnement de la série *Timbre*, en commençant par la bande-image ; ensuite, nous parlerons de la bande-son, pour finir avec une réflexion sur la combinaison audiovisuelle et l'expérience d'audiovision particulière qu'elle rend possible.

### **La bande-image**

Dans ces films de Mirai Mizue chaque timbre est représenté par une configuration visuelle (une « entité ») qui combine des formes, des couleurs et des textures, et qui se déplace et se transforme sur l'écran. Dans tous les films chaque « entité » se répète au rythme d'apparition des timbres sonores. Dans certains films il n'y a qu'une entité (et un timbre) avec plusieurs occurrences. D'autres présentent plusieurs entités.

Très souvent ces configurations visuelles ressemblent des entités zoomorphiques : poissons, oiseaux, poulpes, serpents, vers, insectes, ou même des organismes unicellulaires. Les mouvements impriment aussi des « comportements » zoomorphiques. De fait, plusieurs fois ce sont les mouvements qui leur donnent une allure zoomorphique. Pourtant, ces formes et mouvements ne correspondent pas aux timbres qu'ils représentent, au moins dans un sens « référentiel » ou « iconique ».

Pourtant, dans quelques cas, on peut reconnaître par l'écoute les timbres des « instruments » (piano, harmonica, flute, sonnaïlle, clochette), et la correspondance des mouvements des entités se fait avec *les aspects du processus dynamique des sons*. Il y a aussi des entités non figuratives, c'est-à-dire, qui ne ressemblent à aucun animal ou créature, mais qui thématisent les aspects évolutifs du timbre lui-même.

Le tableau 2 présente la liste de ces entités. On indique les entités qui apparaissent dans d'autres films dans un réseau intertextuel. Il est nécessaire de prévenir que plusieurs

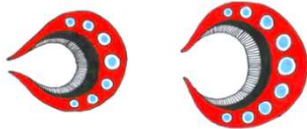
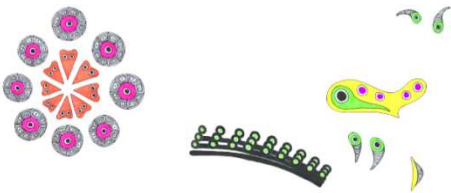
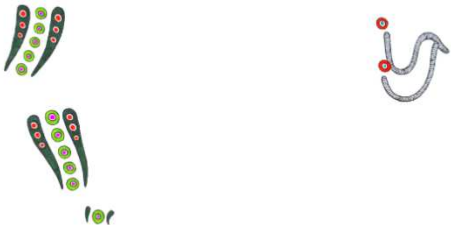
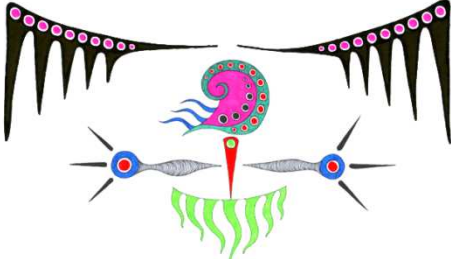
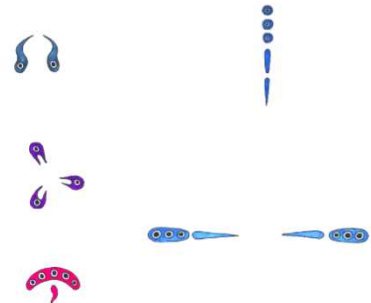
---

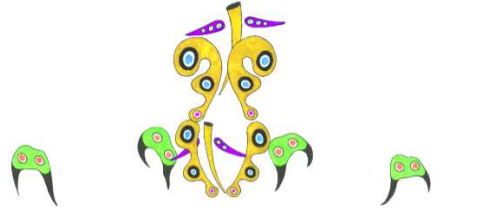

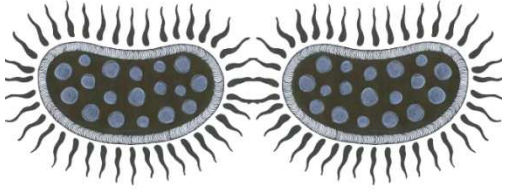
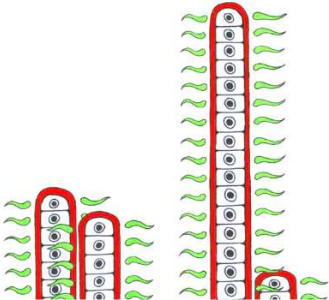
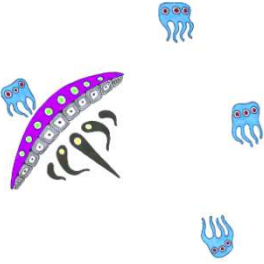
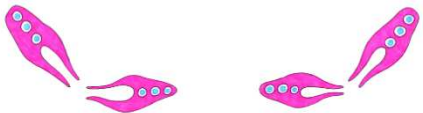

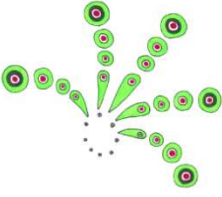

<sup>225</sup> D'après cette description des sons, on peut affirmer que l'opposition ponctuel/duratif peut être toujours relativisée : parce qu'un phénomène sonore de caractère ponctuel peut être décrit comme duratif si la granularité de la description de l'événement sonore analysé est augmentée.

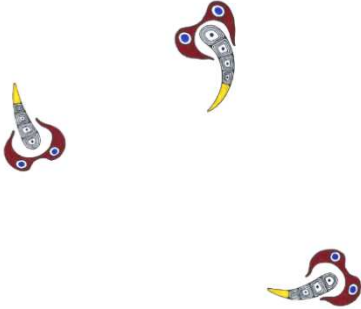
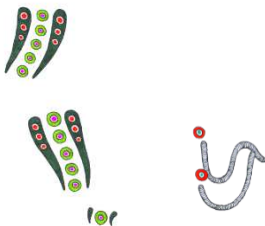






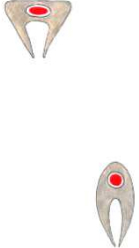
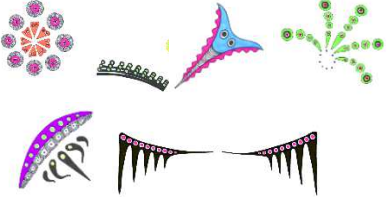
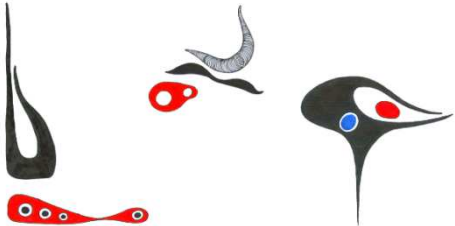
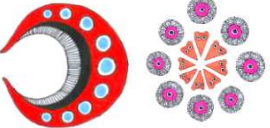


entités ont différents aspects visuels au cours du film où elles apparaissent, et que par conséquent on n'offre ici qu'un de ces aspects, saisi dans un moment de la ligne temporelle.




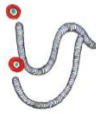
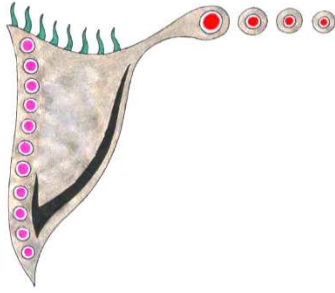
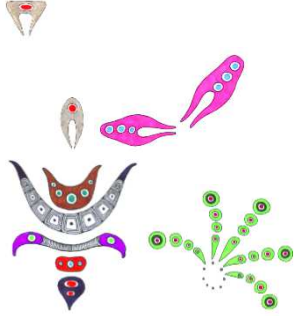
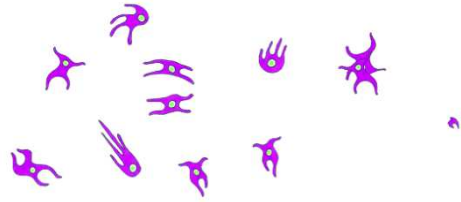


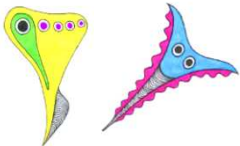
Tableau 5: Les entités du *Timbre A to Z*

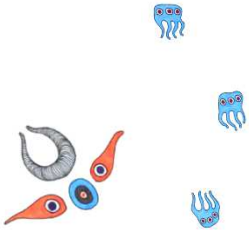
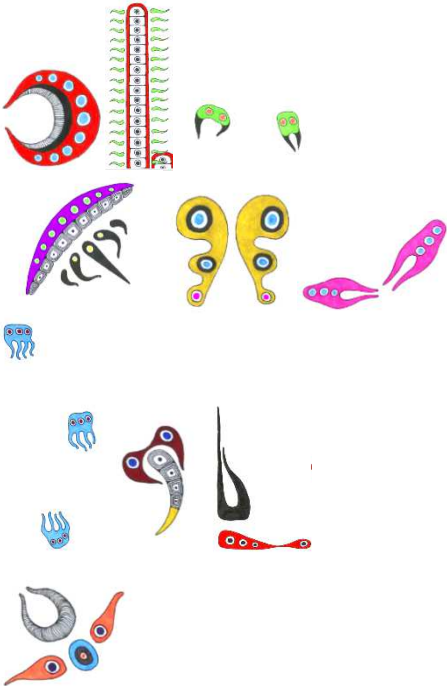
FILM Et # de timbres	Entité « nouvelle »	Entité « réutilisée »
Timbre A 1 timbre		sed diam
<b>Timbre B</b> 3 timbres		
<b>Timbre C</b> 2 timbres		
Timbre D 2 timbres		
Timbre E 6 timbres		

<p>Timbre F 3 timbres</p>		
<p>Timbre G 2 timbres 1 simple 1 composé</p>		
<p>Timbre H 1 timbre composé</p>		
<p>Timbre I 1 timbre composé</p>		
<p>Timbre J 2 timbres</p>		<p>(Timbre F)</p>
<p>Timbre K 2 timbres</p>		
<p>Timbre L 4 timbres</p>		 <p>(Timbre G et B)</p>

<p>Timbre M 3 timbres</p>		 <p>(Timbre C)</p>
<p>Timbre N 3 timbres 1 composé</p>		 <p>(Timbre B)</p>
<p>Timbre O 4 timbres</p>		 <p>(Timbre G)</p>
<p>Timbre P 5 timbres</p>		 <p>(Timbre J et F)</p>
<p>Timbre Q 7 timbres</p>		 <p>(Timbres B, D, G, L et J)</p>
<p>Timbre R 3 timbres</p>		
<p>Timbre S 3 timbres</p>		



		 <p>(Timbres A, B et F)</p>
Timbre T 1 timbre		
Timbre U 2 timbres		 <p>(Timbre C)</p>
Timbre V 5 timbres		 <p>(Timbres K, L, O et Q)</p>
Timbre W 1 timbre		
Timbre X 6 timbres		 <p>(Timbres F, I, K, P et V)</p>
Timbre Y 4 timbres		

		 <p>(Timbres B, G, O et J)</p>
Timbre Z 10 timbres		 <p>(Timbres A, F, I, J, K, M, O et R)</p>

Comme on peut voir dans le tableau, dans les premiers dix films tous les timbres et les entités visuelles sont nouveaux, tandis que, à partir du film onze (Timbre K) quelques entités des premiers dix films commencent à réapparaître dans d'autres films. Les entités peuvent être simples<sup>226</sup> ou composées comme le montre l'exemple suivant :



Figure 36 : Entité simple (du Timbre A)

Source : [Mizue, 2011]

<sup>226</sup> Il n'y a pas vraiment d'entités *simples* chez Mizue puisque, comme on peut le voir dans la figure 4, il s'agit toujours d'une multiplicité de formes assemblées (courbe rouge, points bleus de contour blanc, etc.). On doit parler plutôt de formes *unitaires*. Nous évoquerons plus bas ce caractère multiple des entités par rapport au caractère événementiel des timbres.



Figure 37 : Entité composée (du Timbre O)

Source : [Mizue, 2011]

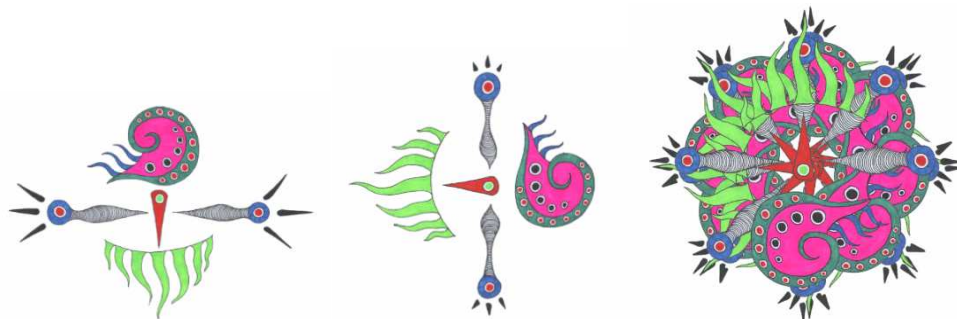
Toutes les entités sont dotées d'un mouvement caractéristique et d'un « arc de transformation (T) » constants, et suivent les mêmes cycles : apparition (A), déplacement (M), disparition (D), bien qu'avec des durées différentes ; c'est-à-dire, la version visuelle d'inchoativité, durativité, terminativité. On peut représenter ce processus comme suit :

$$A \rightarrow (M + T) \rightarrow D$$

Dans quelques cas, il n'y a pas de transformation, mais il y a un autre mouvement typique de l'entité (expansions ou concentrations des sous-entités, mouvements zoomorphiques, etc.). La représentation sera donc la suivante :

$$A \rightarrow (M_1 + M_2) \rightarrow D$$

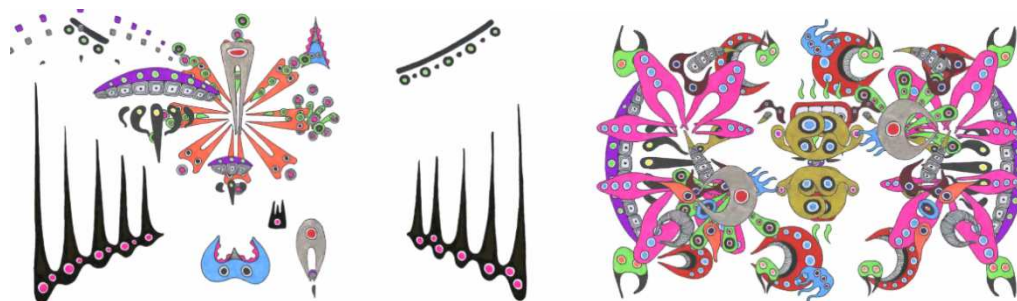
Chaque entité/timbre est ainsi un « paquet » discret qui constitue un processus constant, qui peut être répété. Ce qui peut être variable c'est l'endroit d'apparition de l'entité et l'orientation de son déplacement, variations dues à la rotation de l'entité sur un axe central (et aussi à l'endroit d'apparition). Par exemple, l'entité composée du Timbre D peut se présenter simultanément comme suit :



Figures 38 et 39 : Première apparition de l'entité, Rotation 90°, Rotation (45°), entité répétée et superposée

Source : [Mizue, 2011]

La répétition donne lieu ainsi à une complexité croissante qui varie d'un film à l'autre. En ce sens, les cas extrêmes sont les films *Timbre K* et *Timbre Z*, avec une progression croissante de figures (et de sons, bien sûr) :



Figures 40 et 41 : *Timbre K* et *Timbre Z*

Source : [Mizue, 2011]

Du fait que presque toutes les entités des films *Timbre* sont constituées par plusieurs sous-formes (elles sont ce que le Groupe  $\mu$  appelle des « archi-formes » (Groupe  $\mu$ , 1992), ou d'autres, des « motifs formels » (Barrès, 2006, p. 118)), leur superposition les rend indiscernables après l'accumulation. Quand ils sont répétés, les motifs formels donnent lieu à des « rythmes formels ».

En général, dans ces entités on trouve presque toutes les opérations plastiques mentionnées par Barrès dans le tableau A en ce qui concerne l'analyse image par image : répétition, multiplication, combinaison, élision, extension (Barrès, 2006, p. 118) ; on trouve aussi plusieurs des opérations pour le flux d'images, et des organisations plastiques locales (Ibid.), notamment les « rythmes formels », statiques et en mouvement. Les éléments plastiques qui constituent les entités de Mizue sont principalement des lignes et des formes. La technique plastique de Mizue est fondamentalement le dessin, à partir de l'usage de feutres et marqueurs de couleur sur papier. Les lignes de contour sont toujours noires. Les surfaces des formes sont homogènes et presque sans texture (quelques fois, on note la texture des traces de marqueur). Les caractéristiques du tableau B de Barrès sont donc moins communes chez Mizue<sup>227</sup>.

<sup>227</sup> Les dessins sont scannés et mis en mouvement à l'ordinateur. Ici on peut regarder un résumé du procès d'animation de Mirai Mizue, bien qu'il ne s'agisse pas de réalisation des films *Timbre* : <https://vimeo.com/20222436>. Et une version plus longue (en « temps réel », mais dans ce film, il n'y a que le procès du dessin) : <https://vimeo.com/31382034>.



Un trait particulier dans ces films est le caractère du *fond*, qui renforce la nature plastique de ces animations : il s'agit toujours d'un fond complètement blanc qui établit un espace abstrait, neutre, sans profondeur, bien que les mouvements et les homothéties dynamiques produisent un certain effet de profondeur. La blancheur globale de l'arrière-plan nous fait oublier le fait qu'il s'agit de dessins. On peut contraster cet effet avec l'apparence du fond dans le film *A long day of Timbre* (et aussi dans *Tatamp*), où le fond montre la texture ou matérialité du papier (le support), bien qu'il soit construit avec les mêmes dessins :

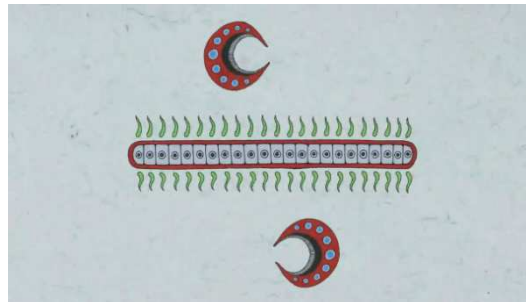


Figure 42 : *A long day of timbre*

Source : [Mizue, 2011]

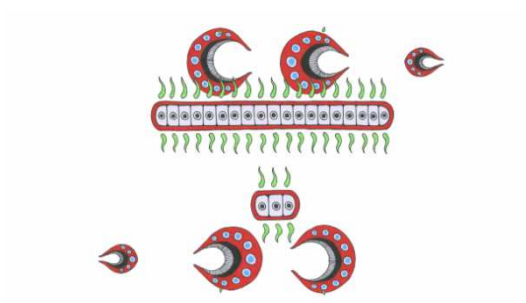


Figure 43 : Timbre Z

Source : [Mizue, 2011]

Il ne faut pas oublier que tous les mouvements et la dynamique visuelle des films *Timbre* sont subordonnés à la structure dynamique et temporelle de la bande-son, comme on montrera ensuite.

D'après Fontanille et Zilberberg, les catégories aspectuelles les plus générales sont : une aspectualité « sautant » qui démarque des limites, et une aspectualité « passante », qui a à voir avec l'établissement de seuils et degrés (Fontanille & Zilberberg, *Tensión y significación*, 2004, pp. 334-335). Dans le même sens Fontanille affirme :

« *L'aspectualisation se présente intuitivement à la fois comme segmentation du procès en aspects et comme modulation continue de ce même procès, et cette synthèse du continu et du discontinu est sans doute un de ses traits spécifiques* ».  
(FONTANILLE, 1991, p. 128)

D'autre part, Greimas et Fontanille introduisent la distinction faite par Michel Colin entre

- Le continu : susceptible seulement de variations d'équilibres et de modulations.
- Le discontinu : comportant ruptures et fractures à raison d'une seule frontière à chaque interruption.
- Le discret : qui requiert deux frontières à chaque interruption (GREIMAS & FONTANILLE, 1991).

Dans la bande-image des films de la série *Timbre* on peut identifier chaque processus des entités comme des modulations continues internes. Chaque processus devient à son tour une unité discrète lorsqu'elle est répétée à des intervalles de séparation qui permettent d'identifier ses apparitions et disparitions (ses seuils). Et il y a aussi les cas dans lesquels des intervalles se chevauchent en générant un flux constitué par discontinuités. C'est le cas représenté par l'illustration 8.

## **La bande-son**

En ce qui concerne la bande-son, chaque film présente un nombre défini de timbres, des plus simples, avec un seul timbre (A, H, I, T, W), aux plus complexes, avec plusieurs (en ordre ascendant de quantité) : C, D, G, J, K, U = **2** ; B, F, M, N, S, R = **3** ; L, O, Y = **4** ; P, V = **5** ; E, X = **6** ; Q = **7** ; Z = **10**). Le fait que le premier film n'a qu'un timbre et que le dernier en a dix peut faire penser qu'il s'agit d'une progression, mais, comme on peut le noter dans l'énumération précédente, ce n'est pas vraiment le cas dans le détail, sinon globalement.

Chaque timbre se répète avec différents intervalles temporels, selon sa durée. Ainsi, les timbres plus courts se répètent chaque seconde (c'est le cas plus général), tandis que les intervalles entre les timbres plus longs sont variables. Mais il y a aussi des cas où des timbres longs se répètent dans des intervalles courts. Ce sont les cas de « chevauchements » entre répétitions d'un même timbre (campanelles, harpes).

Les timbres de Mirai Mizue suivent ainsi la structure aspectuelle décrite par Pierre Schaeffer. On trouve :

- Des timbres courts (percussifs) : le film *Timbre E*, par exemple, n'a que ce type de sons, et aussi le Timbre Q.
- Des timbres de moyenne durée : par exemple J<sub>1</sub>, L ou R.
- Des timbres de longue durée : comme le timbre A, H, I, K<sub>1</sub> ou M.

Le premier type est le plus commun. D'où le fait que l'importance des « attaques » est évidente dans ces films, en consonance avec la proposition de Schaeffer. Cette tendance produit un phénomène particulier : les timbres tendent à devenir percussifs, même s'il s'agit des timbres « mélodiques ». Cela permet la production de rythmes simples à partir de la répétition des timbres courts.

Mais il y a aussi des timbres de durée moyenne et longue qui offrent des complexités intéressantes. Les plus frappants sont les Timbres P<sub>2</sub> et V<sub>1</sub> (et plus fortement, L<sub>1</sub>) puisqu'ils présentent des *transformations internes du timbre*. C'est comme si dans un timbre il y avait plusieurs timbres, ou comme si un même timbre souffrait une série de transformations, en conservant comme sa seule identité la continuité temporelle. Dans ces deux cas, les entités visuelles expriment de façon effective ce phénomène sonore. Regardons alors trois photogrammes de chaque transformation :

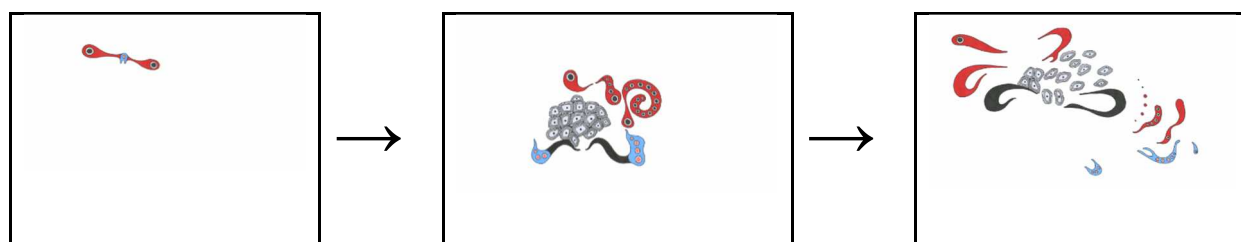


Figure 44 : Photogrammes de Timbre P

Source : [Mizue, 2011]

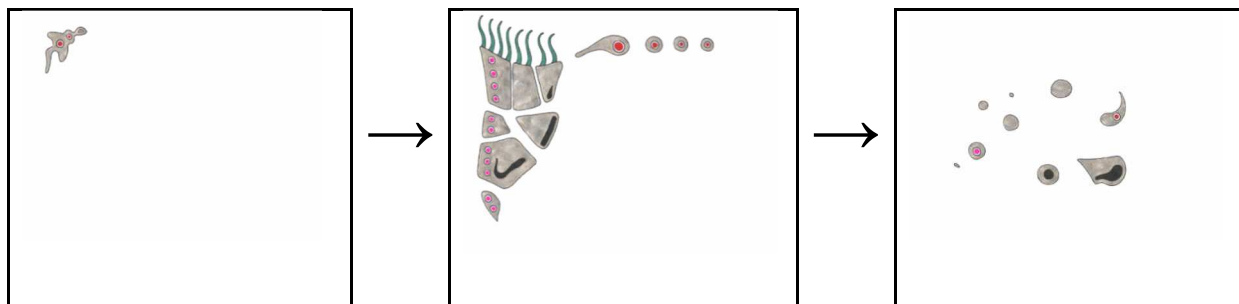


Figure 45 : Photogrammes de Timbre V

Source : [Mizue, 2011]

Il y a aussi des timbres composés : le timbre du film I, qui a un son d'horloge ou chronomètre et un son de flûte, ou le timbre G<sub>1</sub>, N<sub>1</sub> et O<sub>1</sub> (clochette + sifflet). On identifie aussi des timbres superposés, comme le Timbre H, qui commence avec un seul timbre et finit avec deux, superposés. Il y a aussi un autre cas, celui des timbres avec réverbération, comme le timbre J<sub>1</sub>.

### Rythme, arythmie, a-musicalité

Les films de la série *Timbre* sont structurés fondamentalement à partir de la répétition des sons. Dans ce sens, plusieurs des structures temporelles de chaque film peuvent être décrites comme *rythmes*. La définition de « rythme » change selon les approches et les auteurs. Pour Louis Hébert, par exemple, « le rythme peut notamment être défini comme la configuration particulière que constituent au moins deux unités, de « valeur » identique (A, A) ou différente (A, B), dans au moins deux positions se succédant dans le temps » (Hébert, 2016). Pour le Groupe  $\mu$ , cependant, pour qu'il y ait rythme, la présence d'au moins trois unités répétées est nécessaire (GROUPE. $\mu$ , 1992). Selon ces deux perspectives, les sons et les entités de Mizue seront des rythmes (sonores et visuels).

On trouve une description plus précise chez Roger Odin, pour qui le rythme est associé avec le synchronisme :

*...le code du synchronisme intervient chaque fois qu'est produit un message composé par de deux (ou plus de deux) séries temporelles : images + musique, images + paroles, images + bruits, et même images + images (...) Définir la notion de synchronisme mobilise à la fois des relations syntagmatiques de succession et des relations syntagmatiques de simultanéité. Deux « suites » seront dites synchrones si elles sont construites comme ayant le même « rythme » (ODIN, 1990, p. 242).*

Odin prend la définition de rythme de l'Encyclopaedia *Universalis* :

*Le « rythme est organisation des durées, dont il règle la proportion, l'espacement, les groupements ; il est distribution des accents ; il gouverne le rapport respectif » de présence et d'absence des éléments constitutifs de la strate considérée (Ibid., p. 242).*

Mais cet auteur souligne aussi que l'identité de rythme n'est pas suffisante pour produire l'effet de synchronisme, parce que « deux suites peuvent avoir le même rythme tout en étant décalées par rapport à l'axe temporel de référence ». Pour que cet effet de synchronisme se produise, ajoute donc Odin,

*Les deux suites devront répondre aux conditions syntagmatiques suivantes : être construites sur le même rythme (problème de relations syntagmatiques de succession), et être construites de telle sorte que les « seuils » rythmiques coïncident sur l'axe temporel de référence (problème de relations syntagmatiques de simultanéité). Par « seuil » nous entendons les débuts et les fins de séquences rythmiques. (ODIN, 1990, p. 243)*

Selon cette idée, 4 combinaisons sont possibles entre deux « suites » :

1. *Synchronisme*. Même rythme et coïncidence des seuils.
2. *Désynchronisme*. Le rythme est le même dans les deux suites, mais les seuils diffèrent.
3. *Non-synchronisme*. Le rythme et les seuils des deux suites diffèrent.
4. *A-rythmie*. Les seuils de deux suites sont identiques, mais les rythmes diffèrent (les constantes de début et de fin rendent sensibles les différences de rythmes. Les arhythmies peuvent être *disjonctives* –le cas général- ou *contrastives* quand les deux séries ont des rythmes opposés) (ODIN, 1990, pp. 243-244).

Selon cette taxonomie, les timbres de Mizue présentent toujours du synchronisme entre les timbres et les entités visuelles (image + son)<sup>228</sup>. Néanmoins, les intervalles de ces rythmes ne sont pas toujours constants, mais ils peuvent varier dans la ligne temporelle de chaque film, en s'accéléralant ou en se ralentissant. En plus, la présence des différentes

---

<sup>228</sup> Cela donnera l'effet que Michel Chion appelle la « synchrèse ». On reviendra sur ce sujet après.

« suites » entité+timbre dans les films donne lieu aux autres effets décrits par Odin: Désynchronismes, non-synchronismes et arhythmies.

De fait, la présence de sons maintenus et en chevauchement, ainsi que de sons mélodiques (notamment, les harpes et harmonicas) et l'accumulation de timbres de durées différentes, produisent d'autres effets. Ainsi, un film peut commencer de façon très rythmique et finir en un complet chaos, en donnant lieu à l'effet de pur *bruit*. Et même quand il y a une régularité des rythmes, dans quelques films il manque une coordination générale entre ces rythmes pour produire l'effet de régularité globale qui donne le caractère *rythmique* à l'ensemble de phénomènes audiovisuels dans le film. Cela produit un sens d'imperfectivité : c'est comme si tous les films n'avaient pas été finis, mais comme si certains avaient été suspendus.

Il semble que Mizue fait des efforts pour éviter l'émergence de toute structure perçue comme musicale dans ses films. Cet effet de « non-musicalité » ou *d'a-musicalité* se renforce par le fait que les timbres eux-mêmes sont « synthétiques » et non acoustiques, ce qui produit une certaine raréfaction qui nous éloigne de leurs identités originales. De fait, plusieurs des timbres qu'on écoute dans la série *Timbre* ne sont pas reconnaissables comme des timbres d'*instruments*, mais pas plus comme des sons produits par des sources concrètes. Ce sont des timbres *électroniques*, provenant d'un clavier ou d'une banque de sons. De cette manière-là, Mizue induit chez l'auditeur une *écoute réduite*<sup>229</sup>, selon la classification de Schaeffer : il nous oblige à nous concentrer sur les sons eux-mêmes et sur leurs propriétés et non pas sur des objets ou les sources qui les produisent. Le caractère arbitraire des entités visuelles qui leur sont associées renforce cet effet.

À ce sujet il est intéressant de comparer les films de la série *Timbre* avec le film *Tatamp* (Mizue et Suda, 2011) déjà mentionné. Dans ce film, Mizue inclut tout le répertoire de timbres développés dans la série, selon une structure cumulative qui va de l'apparition

---

<sup>229</sup> Définie ainsi par Michel Chion : « L'écoute réduite est l'attitude d'écoute qui consiste à écouter le son pour lui-même, comme objet sonore, en faisant abstraction de sa provenance réelle ou supposée, et du sens dont il peut être porteur. Plus précisément, elle consiste à inverser cette double curiosité pour les causes et le sens (qui traite le son comme un intermédiaire vers d'autres objets visés à travers lui) pour la retourner sur le son lui-même. C'est l'événement que l'objet sonore est en lui-même (et non auquel il renvoie), ce sont les valeurs qu'il porte en lui-même (et non dont il est le support) que vise, dans l'écoute réduite, notre intention d'écoute » (Chion, 1983, p. 33).

d'un timbre ou deux, à la superposition de tous les timbres en même temps, dans la limite de trois minutes et 23 secondes, avec une accumulation chaotique des entités visuelles à l'écran (Voir Figure 40).



Figure 46 : Photogramme de *Tatamp*

Source : [Mizue et Soda, 2011]

Après un silence de trois secondes, les timbres commencent à apparaître à nouveau, mais cette fois avec une structure mélodique et rythmique ajoutée, de sorte qu'ils s'intègrent dans une composition clairement musicale<sup>230</sup>. Pour cette partie Mizue et Suda additionnent d'autres entités qui n'existaient pas dans la série *Timbre*. En particulier, une grande figure de forme circulaire qui semble une horloge ou un métronome, dans une claire allusion de la temporalité régulière de la musique (Voir Figure 41).



Figure 47 : Photogrammes de *Tatamp*

Source : [Mizue et Soda, 2011]

Les films de la série *Timbre*, au contraire, ne parviennent jamais à constituer une structure qui puisse être perçue comme musicale. Au contraire, ils semblent un catalogue

---

<sup>230</sup> Composée par le groupe *Twoth*.

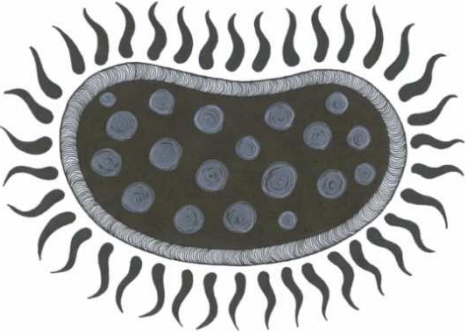
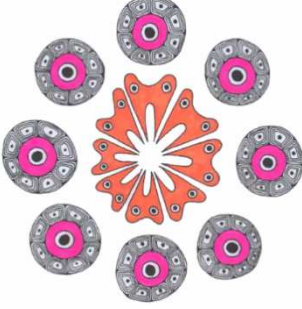
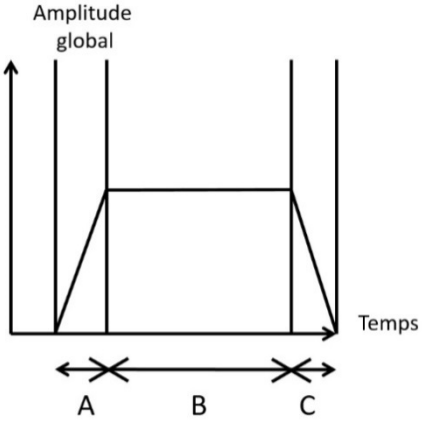
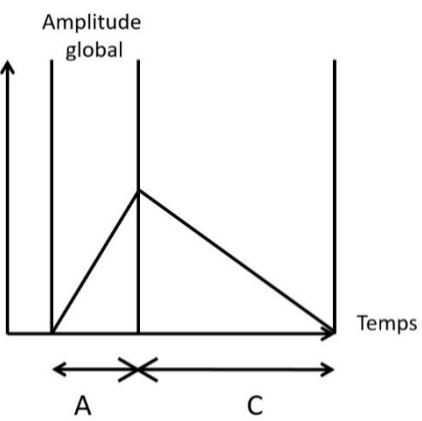


d'expériences sonores, un répertoire destiné à être utilisé dans de futures compositions (comme effectivement le sont *A long day of timbre* et *Tatamp*).

### La combinaison audiovisuelle

On a déjà avancé quelques idées à cet égard, que l'on peut maintenant recueillir ici. La première c'est que les entités visuelles n'ont pas de relation qu'on pourrait dire « motivée » avec les timbres, et qu'il s'agit seulement d'une relation de synchronisation temporelle. Par contre, la structure aspectuelle de la bande-son et de la bande-image est strictement isomorphe (Tableau 6).

Tableau 6: Correspondances audiovisuelles avec des timbres maintenus et avec timbres percussifs

 <p>(Timbre H)</p>	 <p>(Timbre B<sub>3</sub>)</p>
 <p>Amplitude global</p> <p>Temps</p> <p>A B C</p>	 <p>Amplitude global</p> <p>Temps</p> <p>A C</p>
<p>A → (M + T) → D</p>	<p>A → (M + T) → D</p>

Cet isomorphisme profite du phénomène appelé par Michel Chion *synchrèse* :

*Tel est en effet le nom forgé que nous avons donné à un phénomène psychophysiologique spontané et réflexe, universel, dépendant de nos connexions nerveuses, et qui consiste à*

*percevoir comme un seul et même phénomène se manifestant à la fois visuellement et acoustiquement la concomitance d'un événement sonore ponctuel et d'un événement visuel ponctuel, dès l'instant où ceux-ci se produisent simultanément.* (CHION, 2013, p. 191).

Par concomitance Chion comprend « la situation de perception simultanée de sons et d'images, qui donne lieu obligatoirement à des effets calculés, ou non » (Ibid., 190).

Ce phénomène est largement exploité par Mizue, qui parvient à nous faire percevoir les timbres et les entités comme *un seul et même phénomène audiovisuel*. Pourtant, il s'agit d'un type de synchrèse particulière, dans la mesure où elle intègre un *objet sonore*<sup>231</sup> avec une *entité plastique*. Je propose d'appeler ce type *synchrèse figurale*, puisque le figural, tel qu'il est compris par Claude Zilberberg, est un type de figuration présupposée, constante et spécifiée, une sorte de structure profonde qui peut être remplie par des valeurs différentes. Il ne s'agit donc pas d'une synchronisation « iconisante » puisqu'elle ne cherche pas à produire un effet référentiel en associant un son à une source, ou un timbre à un instrument, à un objet ou à un « personnage » reconnaissable. La *synchrèse figurale* peut donner lieu à des « créatures éphémères » dans une zoomorphisation des entités et des timbres, mais cet effet est partiel et il est rapidement suivi par la disparition ou la désintégration des créatures en formes dispersées, en pures vibrations audiovisuelles.

Je voudrais faire deux remarques sur la proposition audiovisuelle de Mizue dans ces films. La première à avoir avec le fait que le timbre n'est pas une propriété des événements sonores, mais un événement en lui-même, avec sa propre structure aspectuelle. Le fait que Mizue ait choisi des entités unitaires, simples ou complexes mais toujours multiples, pour représenter visuellement les timbres est une réussite esthétique : cela nous donne la sensation d'être face à des processus complexes.

La seconde remarque vise une autre distinction aspectuelle : l'axe sémantique ponctualité / durativité. On a déjà vu qu'il y a trois types de timbres dans cette série de films selon leurs extensions. Les timbres « courts », comme ceux qui sont présents dans le film *Timbre E*, peuvent être décrits comme « ponctuels », tandis que les timbres de moyenne durée

---

<sup>231</sup> Entendu dans le sens de Schaeffer comme « tout phénomène et événement sonore perçu comme un ensemble, comme un tout cohérent, et entendu dans une écoute réduite qui le vise pour lui-même, indépendamment de sa provenance ou de sa signification » (Chion, Guide des objets sonores. Pierre Schaeffer et la recherche musicale, 1983, p. 34).

et les timbres longs peuvent être décrits comme « duratifs ». Or la figure 4 nous laisse percevoir que les sons percussifs ont aussi un déploiement temporel, une « micro-durée ». De nouveau Mizue parvient à cet effet en représentant ces timbres à travers une séquence visuelle durative, jamais avec une simple intermittence. En effet, la durativité d'un événement est une fonction de la granularité (l'échelle) dans laquelle on observe un phénomène. Dans ce sens on peut affirmer que Mizue nous offre une sorte de « microscope temporel » (c'est-à-dire, qui amplifie non pas l'espace mais le temps) sans appeler à la rescousse la solution plus banale du ralentissement.

À ce point, je voudrais revenir sur la question de la narrativité et ses rapports avec l'aspectualité. La structure aspectuelle des sons, telle qu'elle vient d'être décrite, nous met face à une expérience perceptive unique de la temporalité, qui la place au cœur de toute aspectualité, puisque les phénomènes sonores deviennent des « objets », « acteurs », figures ou même indices d'emplacements, ou de localisations temporaires, par périodes limitées et avec des étapes d'apparition, de maintien et de disparition. Le monde des *êtres* acoustiques est plus un monde de *devenirs*. En ce sens, les expériences audiovisuelles de Mirai Mizue dans sa série *Timbre* constituent un laboratoire phénoménologique qui nous montre un paysage tensif, peuplé par des processus quasi narratifs ou mieux, pré-narratifs, là où les modulations du changement précèdent les déroulements des actions orientées.

Il faut donc revenir sur le rôle de l'observateur<sup>232</sup> dans l'aspectualisation, quand il s'agit des productions audiovisuelles lues comme non-narratives. Comme cela a été remarqué par la théorie sémiotique (GREIMAS & FONTANILLE, 1991) ; (FONTANILLE, 1991), l'aspectualisation est directement liée avec les processus d'embrayage/débrayage, qui est la dimension déterminante. Dans le cas présent on a vu que c'est la dimension sonore qui donne la structure aspectuelle à l'expérience d'audiovision. D'après Jacques Fontanille, « ...pour être énoncés, un procès ou une figure doivent obligatoirement être ou bien focalisés, ou bien aspectualisés ; s'ils sont aspectualisés, c'est qu'on a fait le choix de manifester telle ou telle composante dans la perspective (et dans l'attraction) du tout » (FONTANILLE, 1991). Dans *Timbre A to Z*, la focalisation est réduite (à travers la neutralisation du fond, le manque de

---

<sup>232</sup> On peut toujours dénoncer l'oculocentrisme de ce concept. Or, d'une façon plus conciliante, on peut l'assumer simplement comme une synecdoque référentielle qui remplace la structure perceptive dans sa totalité multimodale par le sens de la vue. Néanmoins, un peu d'« auriculocentrisme » ne ferait pas de mal...

profondeur, l'annulation des sources sonores, l'appariement audio-visuel) de telle sorte qu'il y a un embrayage presque total, comme si l'audio-spectateur était directement face à un monde possible, comme dans l'expérience du monde naturel.

On peut se souvenir ici de notre petit détour cognitif, pour remarquer l'alignement de l'expérience d'audiovision avec l'expérience générale du monde, où ce sont les aspects dynamiques des mouvements et des transformations qui déterminent la reconnaissance primitive d'événements, et non les catégorisations basées sur les actions et les objectifs. L'aspectualité nous rapproche donc de cet horizon abstrait et ouvert des « préconditions de la signification ». Mais pour que cela soit le cas, il est nécessaire de prendre en considération l'autre dimension constitutive de la production du sens.

### **Rythme, tempo, intensité : la structure tensive de la série *Timbre***

Jusqu'alors, je me suis concentré sur l'aspectualité temporelle de la série *Timbre*. Cela veut dire qu'on n'a rendu compte que du plan de l'expression de la série *Timbre*. Pourtant, j'ai proposé l'hypothèse que la *synchrèse figurale* (la combinaison entre des événements sonores et des entités plastiques) nous met face à une dimension perceptive pure, figurale, qui peut être remplie par des contenus somatiques et des esthésies diverses. Mais pour ne pas laisser complètement ouverte cette question, je voudrais maintenant compléter la lecture de *Timbre A to Z* avec les autres dimensions de la structure tensive proposée par la sémiotique (post)greimassienne.

Effectivement, l'aspectualité temporelle ne nous donne que la dimension extensive de cette structure : on peut mesurer les durées, calculer les intervalles et les positions. Mais les timbres-entités de Mirai Mizue, comme présences qui entrent et sortent des horizons visuels et sonores sont aussi des forces intensives, que produisent chez le spectateur effets de sens différents.

À ce point on peut se souvenir de ce que Roger Odin appelle la *mise en phase* : « le processus qui me conduit à *vibrer au rythme* de ce que le film me donne à voir et à entendre » (ODIN, 2000, p. 38). Odin utilise ce concept surtout pour parler de la narration et de la fiction, en proposant une *mise en phase narrative* et une *mise en phase énonciative*. Or la mise en phase peut donner lieu aussi à un mode de « lecture » des films nommé *mode*

*dynamique* (BUCKLAND, 2004, pp. 89 ; 104-107), ou *mode énergétique* (ODIN, 2000, p. 160-162 ; ODIN, 2011, p. 50-53).

Dans *De la Fiction*, le mode énergétique se confond avec le mode spectacularisant. Si l'exemple qu'Odin analyse en profondeur est la version du film *Metropolis* (Fritz Lang, 1927) fait par Giorgio Moroder (1984), ses derniers exemples sont des films comme *Star Wars*, *Terminator*, etc. Le type de films que Laurent Jullier appelle *de feu d'artifice*. Dans *Les espaces de communication*, par contre, Odin fait la distinction (bien que peu claire) entre le mode spectacularisant et le mode énergétique, en définissant le dernier comme suit:

#### *Définition du mode énergétique*

- *Niveau de l'espace: construction d'un espace image-son (vs un monde ou un espace spectaculaire).*
- *Niveau discursif: blocage (partiel) de la production discursive au profit d'une construction fondée sur les variations rythmiques et les variations d'intensité.*
- *Niveau affectif: relations fondées sur des effets plus que sur des affects*<sup>233</sup>.
- *Niveau énonciatif: construction d'un énonciateur fictif.* (ODIN, 2011, p. 53)

Même si ces critères peuvent être appliqués aux films de Mizue, c'est le troisième critère que je veux discuter. Effectivement, en parlant du vidéo-clip, le type de film qu'active par excellence ce mode de lecture, et du « déplacement dans notre espace culturel de la hiérarchie image/son au profit du son (JULLIER) », Odin affirme que c'est « le signe d'un autre déplacement : celui du sens vers l'énergie » (ODIN, 2000, p. 161). Et dans *Les espaces de communication* : « Il y a toujours production de sens, mais celle-ci passe en quelque sorte au second plan : l'important se joue dans la relation au corps » (ODIN, 2011, p. 52). Mais ici cette relation somatique ne laisse pas d'avoir un certain air péjoratif.

Or on peut prendre un autre chemin pour placer l'énergie dans la dimension sémantique de *Timbre*, et dans l'expérience de son spectateur. C'est le chemin de la sémiotique tensive, qui met au centre ou à la base de la production du sens l'expérience sensible. Dans ce contexte, la description de l'aspectualité temporelle nous a donné la perspective de la *saisie*, relative précisément à « la position, l'étendue et la quantité » (FONTANILLE, 2003, p. 71). Ce qui manque donc dans la description de *Timbre* c'est la

---

<sup>233</sup> C'est moi qui souligne.

perspective de la *visée*, relative à l'intensité, la dimension qui permet de décrire les « vibrations » et les affects du spectateur.

Dans la structure tensive, on place l'intensité et l'étendue sur deux axes qui permettent leur description en termes de « gradients ». L'intensité a deux sous-dimensions : le *tempo* et la *tonicité*, tandis que l'étendue relève du temps et de l'espace. On a déjà vu comment les relations de rythme et de synchronisme tendent plus généralement à l'arythmie. Maintenant, on peut revenir sur l'autre effet de sens qui complémente celui de l'arythmie. Les variations de vitesse des timbres, ainsi que les répétitions et les arythmies, produisent une première variation de *tempo* : de l'ordre des régimes musicaux<sup>234</sup>, régis par les rythmes et les synchronismes, on passe à un régime chaotique, produit dans ce cas par l'*accumulation* d'arythmies et de désynchronismes.

La logique de l'*accumulation*<sup>235</sup> règne dans chaque épisode de *Timbre*, mais aussi dans la série comme ensemble, et au cœur de cette logique se place l'intensité. Et cela est vrai tant pour la bande-son que pour la bande-image. L'*accumulation* affecte le *tempo* : bien que tous les timbres se répètent à vitesses régulières, l'*accumulation* produit un effet d'accélération et la *désagrégation* de timbres et entités produit l'effet contraire : le ralentissement. En ce qui concerne la tonicité, l'*accumulation* produit aussi un effet « tonique », qui a comme limite l'*ébranlement*. La *désagrégation*, dans cette dimension, produit un effet « atone » qui conduit au *calme*, avec la limite de l'*hypnose*.

D'après Jacques Fontanille, « si l'intensité affective est celle qui apparaît conjointement à l'actualisation de la valeur, alors toute intensité associée à un contraste perceptif, et participant de l'actualisation d'une valeur discursive, pourra être portée au compte de la passion » (FONTANILLE, 2003, p. 215). Et l'auteur ajoute : « l'intensité passionnelle du discours a pour corrélat phénoménologique la proprioceptivité, la sensibilité

---

<sup>234</sup> Bien sûr il y a des régimes musicaux qui donnent une place importante aux arythmies et aux désynchronisations. Dans le domaine de la musique populaire, on peut mentionner le cas du *Math Rock*, défini précisément par « une métrique atypique » et par « l'utilisation de mesures asymétriques ». En ce sens, quand je parle de la « musicalité » j'invoque ici le « sens commun », une sorte de *frame* ou de « catégorisation stéréotypique ».

<sup>235</sup> En fait, l'*accumulation* peut être conçue comme un trait stylistique chez Mirai Mizue : une logique similaire se déroule dans les films dérivés de la série *Timbre* (comme *Tatamp*), mais aussi dans films comme *Jam* (2009) ou *Wonder* (2014).

du corps propre qui sert de médiateur entre les deux plans de la sémiotique » (FONTANILLE, 2003, p. 215). La série *Timbre A to Z* « fait vibrer au rythme des images et des sons » notre corps à vitesses et intensités variables en un effet complètement différent de ce que produisent les films spectaculaires ou même les clips vidéo, mais selon la même logique énergétique qui donne au son la baguette du sens, mais pour activer une forme d'audiovision où la « vision » est au service d'une écoute réduite, une sorte d'état contemplatif de l'ouïe, ou de la saturation qui conduit à un chaos qui, heureusement, dure toujours moins de trente secondes.

### **Pour finir : la série dans le cadre des pratiques sémiotiques**

La série *Timbre A to Z* peut s'inscrire dans la catégorie des *Formes brèves audiovisuelles* (PERINEAU, 2013 ; PEZZINI, 2013). Néanmoins, il s'agit d'une forme particulière et peu étudiée qui résulte d'une pratique sémiotique et d'une stratégie consistant en la sériation de la création à partir d'une contrainte auto-imposée : faire 30 secondes d'animation par jour pendant 26 jours, et l'émission correspondante. Tout un tour-de-force, répliqué par les spectateurs qui ont suivi les émissions sur le canal de *Vimeo*, mais aussi par ceux qui ont audio-visionné la série à d'autres moments, ou qui l'ont regardée comme un film unique, dans des pratiques de réception différentes. En tout cas, on peut identifier aussi une structure aspectuelle dans la série comme ensemble, aussi bien à partir de son exhibition sur internet pendant 26 jours, que comme les 26 chapitres d'un court métrage regardé dans une séance.

On a déjà souligné que les films *Timbre A* et *Timbre Z* sont contrastés selon l'opposition *simple* (A : un seul timbre/entité visuelle sans superpositions, répétitions espacées) → *complexe* (Z : plusieurs timbres/entités, superpositions des sons et des entités visuelles, répétitions enchevauchées). Bien que cela puisse donner une perception erronée de la progression dans la complexité, parce que film par film la complexité ne suit pas une progression nécessairement ascendante, l'existence de cette perception nous fait penser que les phases inchoative et terminative imposent à la série considérée comme ensemble une condition de « lecture » selon la durativité, renforcée par la linéarité de l'alphabet qui sert à dénommer chacun des courts métrages. Ce schéma ou « *frame* » de l'alphabet fournit aussi un sémantisme particulier à la série : est-ce qu'il existe un langage ou une écriture des timbres ? C'est un langage « sans contenu », ou simplement la preuve que, comme le pensait Hjelmslev, la substance de l'expression et la substance du contenu ne sont qu'une seule et même chose ?



En outre, la structure globale de la série renvoie à l'instance d'énonciation, non seulement par le fait que l'« auteur empirique » offre directement son produit au public, comme il est très commun aujourd'hui sur internet, mais par le fait que, à travers la réutilisation des timbres/entités dans d'autres films, et de leur répétition à l'intérieur de chaque film, le spectateur peut identifier une logique de production et une caractéristique des médias numériques qui a été identifiée et décrite par Lev Manovich dans son livre *Le langage des nouveaux médias : la modularité* :

*Les éléments des médias, que ce soit des images, des sons, des formes ou des comportements, sont représentés sous forme de collections d'échantillons discrets (pixels, polygones, voxels, caractères ou scripts), éléments qui sont regroupés dans des objets à plus grande échelle, mais tout en conservant leur identité séparément. Les objets eux-mêmes peuvent être combinés à leur tour, conduisant même à des objets plus grands, sans perdre leur indépendance. (MANOVICH, 2005, pp. 75-76)<sup>236</sup>*

Effectivement, la technologie numérique permet à Mizue de transformer une animation faite à la main sur papier, image par image, intégrée avec une bande-son représentant une entité/timbre donnée, dans une unité discrète, un « module » qui peut être répété, multiplié et transformé, et aussi réutilisé dans plusieurs films. Ainsi le travail créatif de Mirai Mizue combine des techniques du dessin traditionnel avec des techniques numériques dans un procès créatif toujours ouvert à de nouvelles expériences. En reconnaissant les timbres/entités comme des unités de cette sorte de *langage des timbres* dont on vient de parler, combinées dans des énoncés différents, le spectateur « sent » la présence de celui qui « parle » le langage des timbres, le responsable de son plaisir et sa perplexité.

---

<sup>236</sup> Il y a une traduction de ce livre au français, mais je cite l'édition en espagnol, et je traduis.

## Conclusion

Tout au long des deux analyses ici proposées, j'ai essayé d'aborder le phénomène audiovisuel du point de vue des pratiques sémiotiques. Cela implique de prendre en compte toutes les pratiques engagées dans la production des énoncés audiovisuels, mais aussi dans sa réception et son interprétation. Pour en faire, j'ai fait appel à la notion de dispositif, compris comme la scène de la pratique où les expériences du sens se produisent réellement.

Le concept de dispositif, cependant, a dû être pluralisé, c'est-à-dire, n'être pas compris comme un seul, mais comme de nombreux dispositifs selon les formes et les techniques de production, mais aussi en fonction de la conception de différentes scènes de réception. Dans ce but, j'introduis le concept de *dispositif animatique*, à savoir le dispositif propre du cinéma d'animation, tout différent de ce du cinéma PVR.

Bien que, selon Philip Brophy le concept de dispositif animatique peut être appliqué à tout le cinéma, dans la mesure où il est le résultat de la mise en dynamique de photographies statiques, il y a une différence radicale : dans le film de PVR il y a des objets (y compris le propre caméra) et des êtres qui bougent par eux-mêmes ou par les forces d'autres « êtres », selon un ensemble de dynamiques du « monde naturel ». Ensuite, c'est le dispositif ce qui «gèle» leur mouvement dans les photogrammes, pour le réinitialiser immédiatement comme un dynamisme restauré, mais que tire profit des caractéristiques de la perception humaine (notamment, l'*Effet phi*) et le dynamisme du dispositif lui-même (le mouvement de la bande-image en avant du projecteur) pour fournir un mouvement presque exact à l'originale.

Le cinéma d'animation, par contre, se produit grâce à la capture, image par image, d'objets statiques dont leur position ou leur configuration est changée avant d'être photographiés (c'est le cas du stop-motion et de la pixilation), ou quand il s'agit de dessins, ils sont complètement remplacés par d'autres dessins avec de petits changements, en reconstruisant les étapes d'un mouvement qui est construit artificiellement. Même dans le cas de la rotoscopie, où le dessinateur retrace le mouvement d'une PVR préalable, il y a un processus de transformation visuelle (voir Groupe  $\mu$ , 1992, p. 139 ; Klinkenberg, 1996, p. 386), qu'affecte aussi l'effet global du mouvement. Cette construction artificielle du mouvement est une caractéristique exclusive du dispositif animatique, et elle a conduit aux

animateurs à développer une étude systématique du mouvement (des objets, des êtres ou du point de vue de la caméra) dans toutes leurs phases et processus, et à systématiser le résultat de ces études à travers des techniques et procédures plus ou moins standardisés. Il s'agit de techniques que permettent la *discrétisation du flux temporel*, et la *composition du mouvement* de façon si détaillée comme les techniques de la musique.

Cette caractéristique a permis au film d'animation de trouver des moyens pour synchroniser le son et l'image qui ont devenue une partie immédiate du dispositif lui-même. Bien sûr, la possibilité de synchroniser le son et l'image avec précision n'est pas exclusive du film d'animation, ou pas même des « talkies » en général, mais elle est une capacité humaine fondamentale que, en fait, est la base de la musique elle-même comme pratique anthropologique, et aussi de la pratique de la danse. La source de ces deux pratiques est, bien sûr, notre existence incarnée, notre corps et ses mouvements, ses rythmes, ses déclinaisons (LEMAN, 2008).

Que est la danse, mais une recherche de synchronisation, de dialogue ou d'échange de nos corps avec la musique (et avec d'autres corps), une musique qui à son tour vient du corps, de son sens du rythme et du mouvement? La compréhension du rythme et du synchronisme est une pratique partagée par le cinéma, la danse et, en générale, par toutes les «arts vivants»<sup>237</sup>. La prééminence des Gags d'humeur physique (accidents, chutes, la maladresse, etc.) qui caractérisent le genre burlesque hérité par le dessin animé, aussi montre le caractère fortement corporel du sens narratif et de l'esthétique de l'animation (WELLS, 2011); (NEUPERT, 2011).

De même, l'expérience de la musique instrumentale est aussi une source évidente de notre expérience du synchronisme : quand nous voyons un musicien exécutant son instrument, nous percevons la synchronisation de leurs mouvements avec la musique qui vient de lui. En effet, la performance musicale a été un favori du dessin animé depuis ses origines, comme l'ont souligné Daniel Goldmark, Michel Selvini (SELVINI, 1995) ou Guido Michelone (MICHELONE, 1995), parmi d'autres.

---

<sup>237</sup> En effet, la danse a été représentée par le cinéma d'animation à plusieurs reprises. Voir, par exemple, **Fuente especificada no válida.**, et **Fuente especificada no válida.**

Ainsi, au stade du film « muet », on a cherché la synchronisation avec la musique d'accompagnement, par exemple, mais dans ce cas, il était la musique ou ses interprètes qui cherchaient se « synchroniser » avec l'image. Puis, avec l'avènement du sonore, l'enregistrement visuel de la caméra a trouvé son équivalent dans les microphones d'enregistrement sonore. Et bien que les deux groupes puissent se construire indépendamment (dans le son post-synchronisé) ce dernier pouvait être enregistré comme une référence pour la bande visuelle, afin de réaliser la synchronisation.

Dans le cas des films d'animation, cela a également été possible et, en fait, a fait partie de leur technique de sonorisation de base. Cependant, la grande découverte du premier film sonore animé a été la possibilité de construire deux flux temporaires (ce de l'image et ce du son) simultanément et de manière synchrone, *comme un pur artifice*. Cette découverte a permis l'émergence et la normalisation des pratiques créatives allant du primaire *mickey-mousing*, dont nous avons parlé à plusieurs reprises, aux processus plus complexes et élaborés, comme la synchronisation des différents niveaux de profondeur de l'image avec les éléments de la bande-son (les voix, les bruits et la musique), en le donnant à chaque personnage ou entité dynamique une piste et une profondeur sonore particulière, en combinant tous les éléments dans une conception sonore globale, selon différents critères esthétiques<sup>238</sup>.

De plus, pour parler en termes sémiotiques du dispositif, je propose de faire la distinction entre le dispositif de production et le dispositif de réception, ce qui nous permet également d'intégrer la théorie de l'énonciation dans ce domaine. Du point de vue du dispositif de réception, il est possible de suivre historiquement la réponse qu'ont eu les dessins animés de Disney dans le public de l'époque, et le fait que la procédure ait été copiée par d'autres studios pour trouver de nouveaux usages et de nouvelles esthétiques. Dans cette évolution, il est nécessaire de mettre en évidence la tendance progressive vers le film musicale des studios d'animation (mais surtout de Disney) et l'assimilation du dessin animé comme un type de

---

<sup>238</sup> À cet égard, l'exemple donné par Jon Newsom est très illustratif. Il s'agit de la séquence connue comme *The Flight to Neverland*, du film *Peter Pan* (Walt Disney, 1953). Cette séquence est particulièrement compliquée à cause de l'usage de la caméra multiplane, mais aussi par sa conception sonore (NEWSOM, 1983, pp. 280-283). Pour audio-visioner cette séquence :

<https://www.youtube.com/watch?v=zf00mEe9EOs>

cinéma éminemment pour le jeune public, une tendance qui serait consolidée avec le mariage entre l'animation et de la télévision.

Ainsi, le dispositif animatique a continué en se consolidant dans le domaine du cinéma, en entrant une série de stratégies (dans le sens que le donne Fontanille à ce terme) de distribution et d'exhibition, où le film a été coordonné avec d'autres pratiques de l'industrie de la consommation associées au public d'enfants, mais aussi avec des stratégies de production, comme les adaptations de littérature ou de contes pour enfants. Un élément clé de cette stratégie a été la composition musicale et la commercialisation dans l'industrie de la musique, par le choix des chanteurs célèbres qu'interprètent les thèmes musicaux des films.

Dans sa mutation vers la télévision, le dispositif animatique s'est adapté facilement à la "scène domestique" de réception, étant associé à son tour avec les modes de vie des sociétés industrielles : liée avec les loisirs des enfants, l'animation a commencé à occuper une partie de son «temps libre», en faisant partie de la programmation des weekends, etc. Bien sûr, cet enfantillage n'est qu'une des nombreuses tendances qui a pris le dispositif animatique tout au long de son histoire, bien que l'une des plus répandues, une tendance qu'a finir par assimiler ce dispositif au sein du dispositif cinéma de divertissement traditionnel.

C'est dans ce sens que tous les outils descriptifs du rôle du son au cinéma mentionnés dans le chapitre 1 pour le cinéma PVR peuvent également être appliqués au cinéma d'animation narratif, alors que celui-ci partage de plus en plus les conventions narratives et les stratégies avec celui-là. La tendance vers le réalisme visuel dans l'animation numérique a également renforcé cette assimilation, ainsi comme la tendance à incorporer l'animation dans les effets spéciaux au cinéma PVR.

Cependant, dans cette thèse je choisi de ne pas me concentrer sur le développement du dispositif animatique tel qu'il a évolué dans le cartoon américain et dans la production industrielle. J'ai pris comme sujet une autre tradition du film d'animation qu'a également donné un rôle de premier plan à la bande sonore : le film appelé abstraite ou expérimentale, et en particulier la pratique connue sous le nom de son dessiné, ou son animé. Les différences radicales de ce type d'animation avec le cartoon ou avec l'animation industrielle, soulèvent la

question de savoir si n'est pas nécessaire pluraliser également le concept de dispositif animatique, ou s'il s'agit tout simplement de différentes utilisations du même dispositif<sup>239</sup>.

Une première réponse peut être trouvée dans la petite généalogie que j'ai proposée au chapitre 3, où, d'après Patrick Barrès, j'associais le film expérimental des «artistes sonores» avec les pratiques de l'atelier des arts traditionnels. Toutes les techniques du dessin sonore identifiées dans ce parcours (le son tracé ou rayé directement sur le film ou le son tracé sur papier et photographié dans la bande d'images), rendent la relation de l'animateur avec l'objet-support plus proche de ce de l'artiste plastique avec ces matériaux, et plus éloignée du labour du travailleur spécialisé dans le cinéma industriel, où les travaux liés à la production du son sont très diversifiées. Ainsi, dans le domaine de l'énonciation «en acte» c'est une personne ou une équipe très réduite, ce qui est en charge de la production globale, y compris la bande sonore.

Cela permet que l'animateur apprenne à connaître et à manipuler leur matériel à des degrés de précision de très haute teinte, en développant des codes très élaborés. Ainsi, la pratique du son dessiné le permet à une personne comme McLaren (et avant lui, à Pfenninger) de «descendre» au niveau des unités minimales et des traits des sons, pour créer son propre code d'équivalences avec les formes visuelles, et la création des configurations sonores qui ne peuvent pas être atteints autrement. Dans ce parcours, j'ai suivi le déploiement du dispositif animatique de production, en montrant comment sa composante matérielle a permis aux dessinateurs sonores créer tout un dispositif mental, selon la terminologie Siroi-Trahan.

Du point de vue d'un dispositif de réception, dans l'analyse des *Synchromie* j'ai essayé de montrer comment McLaren a introduit une dimension «didactique» dans sa conception générale, en présentant au début des équivalents son/image d'une façon simple, et en

---

<sup>239</sup> L'existence d'un film comme *Fantasia* (Walt Disney, 1940), qui parvient à intégrer les formes narratives et esthétiques propres du cartoon avec d'autres plus près du film expérimental «synesthésique», ainsi que l'intégration d'une technologie sonore comme le Fantasound, qui a ouvert de nouvelles possibilités d'audio-vision dans les théâtres, nous font penser que cette seconde idée est également plausible, ou du moins que la distinction entre les différents types de dispositifs animatiques est floue.

incrémentant sa complexité tout au long du film. En plus, l'incorporation à l'écran des formes visuelles de la bande sonore offre également au spectateur la possibilité de capturer les équivalences exclusives de cette technique entre l'image et le son. Ainsi, il semble inévitable dans ce genre d'animation, l'incorporation d'une dimension méta-sémiotique, qui permet au spectateur d'acquérir les compétences nécessaires pour interpréter ce type de film.

Cela me conduit à une série de questions : est-il possible d'interpréter correctement une animation comme *Synchronie*, d'une perspective *immanente*? Est-ce que ce *texte filmique* offre lui-même les clés de son interprétation? Pourrait quelqu'un avec l'ignorance totale du « comment a été fait » découvrir la richesse de ses équivalences visuelles-sonores? Pourquoi «sauter» au niveau des pratiques pour son interprétation?

Dans le cas des deux lectures «en profondeur » mentionné dans cette thèse, ceux de Terence Dobson et d'Utako Kurihara, le point de départ semble être toujours la technique avec laquelle le son a été produit, et le fait que «ce que vous voyez est ce que vous entendez ». Cette information ne figure pas dans le film, même pas dans son paratexte, et seulement si l'on a vu des documentaires ou des commentaires de McLaren lui-même, c'est-à-dire, si l'on a eu accès à l'information "externe", on pourrait découvrir que ce qu'on voit dans l'écran est la bande-son. Pour capturer ce sens, donc, il est nécessaire d'accéder à un réseau d'intertextes avec une emphase métatextuelle forte. Ces intertextes ont pour thème central la pratique de création de l'artiste, son matériel et sa technique, y compris le travail sur le support, la construction des codes des équivalences son/image, etc. Autrement dit, les informations générales de ce que j'ai appelé son *dispositif de production*.

Ainsi, une partie importante de l'analyse de la pratique du son animé a consisté en la reconstruction du réseau intertextuel que "entoure" le film, que l'explique et le complémente. Et bien que l'analyse de ces informations ne remplace pas l'analyse directe de la pratique de création « en acte » -difficulté infranchissable dans l'étude des produits historiques-, elle est un complément d'information essentielle qui devient parfois plus précieux que de celle qui fournit le film lui-même. C'est le cas du spectateur occasionnel qui trouve le film "ennuyant" ou " sans sens", mais qui est surpris et heureux quand il est informé des détails de leur production.



Cependant, bien qu'on n'ait jamais accès à ces informations "externes", le film lui-même offre des caractéristiques visuelles et sonores que « sont l' » et qu'un spectateur attentif pourrait saisir et interpréter. Mais l'accès à ces caractéristiques dépend inévitablement du dispositif de réception, notamment de sa dimension matérielle. Par exemple, certaines des caractéristiques sonores manipulées par McLaren avec son système de cartes dépendent d'un système d'amplification spécifique, stéréophonique, qui permet une certaine qualité de la reproduction. Avec un système monophonique seront perdus certains mouvements ou déplacements du son, ni le sens de la profondeur qui construit par moments McLaren dans son court-métrage, et qui dépend de la manipulation de l'intensité et de la réverbération décrites dans chapitre 3. Donc, en ce qui concerne le son et l'expérience d'écoute, où se trouve la limite de leur « immanence »?

Bien sûr, il y a un sens qu'on peut toujours capturer dans les pires conditions de réception audiovisuelle (quand l'on a au moins les signales de l'image et du son), et provient de la synchronisation détaillée entre les figures géométriques (carrés et bars), leurs mouvements et leurs transformations, et les structures rythmiques et mélodiques (leur aspectualité) de la bande sonore. Ce niveau de base est expliqué par l'effet de synchrèse dont nous avons parlé, et se produit presque automatiquement dans le plan purement perceptif. La construction d'un objet de sens à partir du métalangage sémiotique peut être faite avec les outils offerts par cette discipline, y compris le niveau énonciatif, comme dans la lecture présentée ici dans le chapitre 3<sup>240</sup>.

Cependant, il y aura toujours une question ouverte sur la pertinence de l'information perceptive qui est perdue pour ne pas avoir accès aux conditions idéales de perception. Comment affecte-t-elle la lecture de la dimension sémantique profonde l'ignorance de sa profondeur sonore ? Cela me porte à croire que les frontières de la textualité filmiques ne sont pas tout à fait claires, et qu'elles se confondent avec le propre dispositif. Si un cinéaste (ou une équipe de production) a le contrôle sur le choix du système de l'amplification et de la reproduction sonore requis pour saisir pleinement l'expérience audiovisuelle, et il calcule les effets de sens dans ces conditions, où l'on pourrait vraiment tracer un *plan d'immanence*, dans le cas du cinéma et de l'audiovisuel, c'est *dans le dispositif lui-même*. Je veux dire que tout plan d'immanence tracé sur la textualité filmique sera fragmentaire et incomplet, et que

---

<sup>240</sup> Voir aussi l'annexe 1 avec la description par séquences de *Synchronie*.

l'analyste qui fait appel à cette réduction méthodologique de ce type devrait être conscient de ce fait.

L'accent de la sémiotique et de la narratologie audiovisuelle sur le dispositif du film institutionnel classique a fait que tous les aspects de la gestion de la perception opérée par le dispositif deviennent supplémentaires de la composante narrative, à travers lequel on a accès aux autres niveaux sémantiques. D'où que la matière perceptive (sonore et visuelle) est envisagée seulement dans la mesure où elle permet de construire une diégèse, de représenter un point de vue (ou un point d'écoute), et de participer à la distribution générale de l'information narrative et ses modulations; or cette information perceptive est défini comme «excès», comme le font David Bordwell et Kristin Thompson dans sa définition du style filmique, en utilisant l'expression de Roland Barthes pour confiner à ce nom tout ce qui ne peut pas être assimilé dans le style (c'est-à-dire, tout ce qui n'est pas contrôlé).

Ce détournement n'est pas possible dans le cas des productions audiovisuelles comme ceux que j'ai discutés dans cette thèse. Ces films proposent à nouveau la question de ce qu'a été décrit auparavant des sémiotiques monoplanaires, ou des non-langages. En ce sens, Jacques Fontanille invité à effectuer un mouvement théorique qui peut être très utile pour comprendre les phénomènes expressifs dans l'audiovisuel: assumer tous les langages comme non-langages, pour accéder à ce niveau de «sens perceptif» que, dans la formulation de Fontanille, était renvoyé à l'« acte d'énonciation » de la pratique de production de sens (Fontanille, 1995).

Ici, j'ai suivi un chemin similaire, mais pour localiser le niveau de la *praxis énonciative* dans le dispositif, où se trouve enregistrée matériellement la configuration perceptuelle de l'expérience audiovisuelle. S'il y a une idée qui semble pour moi encore fondamentale dans la théorie du dispositif, même dans la formulation originale de Baudry, est l'idée que la technique est porteuse de sens, mais ce sens n'est pas nécessairement une «idéologie» articulée avec une « production de la subjectivité ». Je ne nie pas que, en effet, cela puisse être le cas, mais je pense que ce sens est, en principe, beaucoup plus simple et en même temps, plus profond.

## Le dispositif audiovisuel et l'actant de l'audiovision

J'ai déjà avancé que le sens qu'est construit et que circule dans les dispositifs est, en principe, corporel, comme l'est tout sens que l'homme le donne au monde et à l'expérience, selon l'idée partagée par les théories de la «cognition incarnée» et par la sémiotique tensive. Par conséquent, dans tout dispositif il y a une dimension sémantique qui est produite et maintenue au niveau de la perception, alors qu'elle est déjà configurée sémiotiquement. La particularité du dispositif animatique de ce qu'on a appelé *synchrèse figural* dans l'analyse des films de Mirai Mizue c'est que la production de sens est réduite, ou mieux, est concentrée, dans ce niveau de l'expérience incarnée. Pour les films en général (y compris le film narratif traditionnel d'animation), ce niveau cède la place à d'autres formes de catégorisation sémantique, qui seront «versées» en figures et mises en syntagme dans des configurations narratives, etc.

Je vais donc essayer à formuler la structure élémentaire de la signification à ce niveau, à la suite de certaines des idées contenues dans les chapitres précédents, et d'autres idées provenant des traditions associées au paragraphe précédent avec le sens incarné.

Le dispositif audiovisuel a été construit pour un type de manipulation du spectateur que j'ai appelé *faire-percevoir*. Ainsi, la première opération du dispositif sera la localisation spatiale de l'observateur à partir d'une «logistique de la perception»<sup>241</sup>. Ainsi, ce spectateur abstrait est avant tout un *actant positionnel* un centre déictique, situé dans l'espace géré par le dispositif. Baudry décrit en détail comment opère cet emplacement des spectateurs du cinéma traditionnel, associé à la salle obscure du cinéma, et à l'expérience collective d'un public anonyme. Cependant, cet emplacement peut être effectif dans d'autres conditions matérielles, dont la seule exigence est la présence de sources contrôlées de son et de lumière.

À ce stade, il est important de dire que si le dispositif nécessite l'étatisme du spectateur, est par la configuration perceptive de la vision, parce qu'il est toujours possible d'écouter en mouvement, comme bien le savent les conducteurs de voiture ou les personnes

---

<sup>241</sup> Je prends cette expression de Paul Virilio, mais je l'interprète dans son sens le plus élémentaire comme le contrôle de la perception par le biais d'une médiation technologique. Pour le sens original, voir **Fuente especificada no válida.**

qui font du jogging. Cet étatisme est seulement partielle, et applique uniquement à ce qui Fontanille appelle le « soi », associé au corps propre (FONTANILLE, 2011, p. 13). Parce que l'autre aspect du corps, la *chair*, concernant le « moi » est un centre déictique en *vibration* permanent et, dans le cas du cinéma, cette vibration est suscité par les sons et les images fournis par le dispositif.

En ce sens, l'étatisme est en fait une interruption temporaire de la locomotion, et non, comme pensait Baudry, une suspension de la motricité qu'il associait à la «immaturité moteur» de l'enfant, et puis à l'inhibition de la motricité dans le rêve. Selon les conditions du dispositif de réception, en plus du mouvement des yeux dans l'exploration de l'espace visuel – l'écran dans ce cas- le cou et la tête se déplacent également dans cette exploration, qui est articulé avec la perception des sources son réel de la pièce, qui peuvent être distribuées autour. De même, le corps réagit à des stimuli reçus, même s'eux provient d'un monde artificiel et fictif, comme on peut le vérifier lors d'une séance de cinéma d'horreur: le spectateur peut fermer les yeux, couvrir son visage avec ses mains, tourner la tête dans la direction opposée à l'écran, même sauter après d'une apparition ou un son soudaines.

À cet égard, une alternative à la conception classique du dispositif peut être la théorie du "PECMA flow" de Torben Grodal. PECMA, acronyme anglais de Perception, Emotion, Cognition et Motor Action, est selon cet auteur le flux continu de l'activité cognitive du spectateur audiovisuel, qui commence par la perception des stimuli visuels et auditifs, se poursuit avec des réponses émotionnelles (la première activité cognitive qui est activée par des stimuli), suite dans les processus cognitifs de haut niveau associés à la catégorisation, la formulation d'hypothèses et inférences, etc., et se termine par l'action motrice, également présente dans le cas de l'expérience filmique, bien que contrôlée à partir de la connaissance du caractère fictif ou artificiel de ce qui est vu et entendu (Grodal, 2006).

Dans ce contexte, l'actant positionnel de la scène de réception audiovisuelle participe d'une *visée* et d'une *saisie* (FONTANILLE, 2003, 161). Le premier niveau de l'expérience audiovisuelle, le plus élémentaire et automatique, est-ce de la visée : l'actant de l'audiovision est un « sujet vibratoire », pure chair sensible dans un espace perceptif ouvert. Dans le synchrèse figural cela est le mode en vigueur, et le spectateur est articulé avec sa cible par le rythme et l'intensité de l'expérience, où c'est surtout le son ce qui impose ses structures et ses déploiements. Dans ce cas, si l'actant blanc est une configuration sonore et visuelle, l'actant

de contrôle est opéré par le dispositif lui-même et sa gestion des traits d'accès à cette expérience.

Par exemple, dans les animations de Mirai Mizue qu'on a analysée ici, un composant visuel qui fonctionne comme actant de contrôle c'est le fond blanc et indifférencié, ou peut-être mieux, « l'absence de fond », qui permet le déploiement des entités-timbre, son parcours éphémère et le cyclique de ses apparitions et disparitions. Il s'agit d'un horizon neutre, ce qui permet l'émergence des présences visuelles et sonores. Dans le film de McLaren il y a aussi, en principe, un actant de contrôle neutre : le fond noir, mais il est bientôt remplacé par la grille des barres qui régit la répartition des carrés et des colonnes de couleur, leur organisation en groupes, etc.

Dans la dimension sonore, le contrôle est déterminé par le dispositif de distribution des sons. C'est à ce niveau que les caractéristiques techniques du dispositif de réception sont indiscernables de leurs composantes textuelles. Parce que, soit par le biais d'un système de haut-parleurs ou d'un casque, il est nécessaire dans les deux cas (Mizue et McLaren) que l'information sonore soit distribuée *en stéréo*, pour permettre de percevoir les déplacements du son.

D'autre part, il y a effectivement un composant qui peut être entendu comme contrôle, avec indépendance des conditions techniques du dispositif : le *silence*, cet horizon de l'espace sonore dont nous a parlé Don Ihde, et qui joue ici un rôle actif en tant que facteur de différenciation des présences sonores qui nous donnent les animateurs abstraits (les timbres Mizue, les notes discrètes McLaren).

Dans le dispositif animatique (tant dans le figural des animations abstraites, comme dans le dispositif traditionnel) est fondamentalement dans la *visée* où la production de sens se fonde. Bien sûr, il y a aussi *saisie*, mais cela ne se déploie que quand il s'agit de comprendre la façon dont il opère la « résonance » entre le corps qui perçoit (la chair) et les rythmes, les vitesses, et le tempo de l'expérience activée par les images et les sons.

Maintenant, cette idée de la *visée* comme pure réaction pure, et l'actant comme pur emplacement dans un espace géré et « manipulé » par un dispositif, peut donner une conception passive de la perception, comme l'objet de cette manipulation opérée par le dispositif, ou de cette « logistique de la perception » qui semble être conduite par une volonté

invisible. Afin d'éviter cette interprétation, nous avons décrit la relation de source / cible comme orientée du spectateur vers les stimuli et non à l'inverse, ce qui est toujours possible: en effet, le spectateur a traditionnellement été décrit dans le modèle de communication, comme le « récepteur » plus ou moins passif, d'un « émetteur » qui fournit des informations que celui ne peut qu'accepter ou rejeter.

Les concepts sémiotiques de source et cible sont dépourvus de cette charge, et permettent de positionner ainsi un système de directions orientées selon le critère d'analyse. Ainsi, selon cette deuxième orientation on peut penser que, en effet, le dispositif audiovisuel de réception a été conçu pour positionner le spectateur dans un espace géré, de sorte qu'il reçoit un ensemble de stimuli visuels et sonores provenant des sources contrôlées et avec directions prédéterminées. En ce sens, le spectateur est la cible d'une expérience préconçue pour lui.

Source et cible sont donc organisées selon deux programmes différents: ce d'un énonciateur, matérialisé sous la forme d'un dispositif, et ce d'un énonciataire positionné activement face à une expérience dont il veut trouver un sens. De l'idée de manipulation, nous pouvons passer à l'idée de contrat, ou mieux, de promesse mutuelle<sup>242</sup>: le dispositif délivre des stimuli configurés à travers de quelques objets et supports distribués dans l'espace, et le spectateur est « engagé » à son interprétation et à son « expérience » avec toute la puissance active de sa perception. Mais cela implique que la conception de l'actant de l'audiovision ne peut pas être purement positionnelle, car comme a dit Fontanille, « les actants positionnels ne font rien par eux-mêmes ; ils occupent des places, ils sont mus par une énergie qui les déplace » (FONTANILLE, 2003, p. 163).

Le spectateur audiovisuel, dans la mesure où il est un sujet perceptif actif, il doit être compris aussi comme un actant transformationnel : le sujet d'une audiovision « profonde ».

---

<sup>242</sup> Je remercie Silvie Périneau de me faire voir la subtile différence entre contrat et promesse, dans la séance du séminaire doctoral où j'ai présenté les premiers développements de ce travail.

## Audiovision, visu-audition et écoute profonde

L'activité du spectateur ne va pas "contre" l'emplacement perceptif que le dispositif l'offre. Au contraire, l'appareil attend et favorise sa participation active. À cet égard, il faut souligner que ce que nous avons appelé audio-vision, d'après Michel Chion, n'est pas un "effet" perceptif des médias audiovisuels sur le spectateur, mais une *activité perceptive*. Dans le cas du dispositif que j'ai essayé de décrire ici, cette activité est particulièrement importante, parce que, dans la mesure où dans les films analysés il n'y a pas de récit fort ou des contenus thématiques clairs, c'est dans la *perception elle-même* où le spectateur doit trouver le sens. Cela implique que le dispositif animatique de la synchrèse figural exige au spectateur une *exploration attentive* des deux flux perceptifs dont il est exposé : le flux visuel et le flux sonore.

Michel Chion a fait la distinction entre l'expérience de l'audiovision, le genre de perception "standard", pour ainsi dire, des médias audiovisuels, et la visu-audition :

« Le terme symétrique de *visu-audition* peut s'appliquer à un type de perception audio-visuelle consciemment concentré sur l'auditif (comme dans le cas du concert ou d'une scène de film montrant une exécution musicale, mais aussi lorsqu'on est attentif aux propos tenus par quelqu'un) et où l'audition est accompagnée, renforcée, aidée ou au contraire déformée ou parasitée, mais en tout cas influencée par un contexte visuel qui peut amener à projeter sur elle certaines perceptions » (CHION, 2013, pp. 186-187).

Les productions animées analysées dans cette thèse, choisis précisément en raison de l'importance qu'elles le donnent à la composante sonore non seulement dans sa configuration formelle, mais aussi dans sa thématique, ont besoin d'une sorte d'attitude perceptive plus proche de cette idée de visu-audition. À cet égard, l'accent sur *l'écoute* entendue comme exploration active devient fondamental. Nous avons déjà vu que Michel Chion récupère la classification des types d'écoute proposés par Schaeffer, en les réduisant à trois (écoute causale, l'écoute réduite et l'écoute sémantique) et les applique à l'expérience audiovisuelle. Laurent Jullier a également proposé une classification des écoutes d'après les intérêts du spectateur ou de l'analyste (écoute esthétique, écoute technologique, écoute cynégétique et écoute structurale).



Maintenant, je veux proposer un autre type d'écoute qui, à mon avis, est plus pertinente en relation avec l'expérience suscitée par des films commentés ici. Il s'agit d'une sorte d'écoute qui a été identifiée pour l'expérience musicale, mais qui trouve dans ces films "sonores" (dans la mesure où chez eux le son régit l'image) un sens particulier. Il s'agit de *l'écoute profonde (deep listening)*, une expression utilisée par le compositeur Pauline Oliveros et récupérée en tant que catégorie d'analyse de l'expérience musicale par le philosophe Joel Krueger.

D'après Krueger, l'écoute profonde est un mode volontaire de concentration perceptive soutenue et de sensibilité émotionnelle, qui se concentre sur la dimension spatiale ouverte par la musique. L'expérience musicale est unique, selon Krueger, à cause de la complexité de son contenu *spatial*, parce que l'espace de la musique est *perçu précisément dans sa temporalité*.

Krueger rappelle un passage de Merleau-Ponty (KRUEGER, 2011, p. 69), l'un des rares paragraphes où le philosophe français parle de la musique : en disant qu'un morceau présente une phrase musicale après l'autre, mais l'enchaînement crée une totalité, un monde que n'existe que dans la musique, « dans le quartier de Debussy ou le royaume de Bach ».

Ce monde créé par la musique n'émerge que comme le résultat de l'exploration d'un morceau musical, d'une attention particulière à sa dynamique temporelle, laquelle construit progressivement une topographie sonore, en invitant à continuer l'exploration. Le déploiement temporaire de la musique est le mouvement qui commence à ouvrir son espace sonore interne.

Dans un autre passage, Merleau-Ponty (KRUEGER, 2011, p. 70) établit une distinction entre l'espace de la salle de musique où on écoute un concert, et l'autre espace ouvert par la musique, une nouvelle dimension qui vient de la musique elle-même. Selon Krueger, on peut distinguer entre un espace musical intérieur (espace structurel, interne au morceau lui-même, sa structure syntaxique, constituée de ses composants, tel que des progressions rythmiques, tons, etc.) et un espace externe (l'espace « local », égocentrique, qui permet de placer les sons avec précision dans une salle).

Selon Krueger, dans l'écoute profonde nous « énonçons »\* une fusion expérientielle de ces deux formes de spatialité, de sorte qu'aucune d'elles ne prend une priorité phénoménologique sur l'autre. L'écoute profonde est une forme d'écoute immersive dans laquelle le sujet est sélectivement orienté dans un morceau de musique traitant activement ses différentes caractéristiques sonores et les interrelations entre eux, tout en maintenant simultanément un état de réceptivité émotionnelle. L'écoute profonde est un mode trans-actif d'écoute qui implique des processus tels que l'exploration, la sélection, la modification et la mise au point de l'attention. L'écoute profonde implique le développement d'un état de champ auditif, de sorte que l'auditeur se sent habiter la profondeur du champ sonore en le conduisant à une réaction émotionnelle et affective accrue face à la situation musicale.

Dans le cas de la *visu-audition* de l'animation sonore, l'écoute profonde fonctionne dans la mesure où les combinaisons audiovisuelles attirent notre attention sur les éléments sonores, sa « matérialité » et sa « spatialité ». Le fait que l'image est réduite à des formes en deux dimensions, dans un effet de surface (sans profondeur) tant chez Mizue comme chez McLaren, aide à que la sensation spatiale d' « immersion » soit fournie par le composant sonore à travers des transformations de l'intensité, la hauteur et la réverbération.

Dans le cas de *Synchromie* c'est clair la transition de la simple succession de progressions musicales dans une ligne temporelle à sa spatialisation à travers des échos et des variations de l'intensité (voir annexe 1). Dans *Timbre A to Z*, on arrive à sa dimension spatiale à travers de la superposition de couches sonores que, dans les cas les plus extrêmes (Timbre Z, par exemple) remplit complètement l'espace d'écoute, en le saturant.

---

\* Le verbe « énoncer » (to enact) n'existe pas en français et en anglais a un sens « technique » dans le langage des sciences cognitives, notamment dans la théorie de l'énonciation (*enaction*). Voir **Fuente especificada no válida..** À propos de l'usage du terme, et notamment de son application aux études du son, Renaud Meric nous rappelle que le terme « énoncer » est synonyme de « faire-émerger » et il cite Pierre Lavoie, le traducteur de l'ouvrage *Invitation aux sciences cognitives* de Francisco Varela, en expliquant « le choix et l'« étymologie » du terme « énonciation » et de l'expression « faire-émerger » : « *Le premier de ces deux néologismes [énoncer] reproduit le terme anglais « énonciation », en tentant de préserver la proximité entre « action » et « acteur » ; le second [faire-émerger] nous vient de la tradition de la phénoménologie qui traduit hervorbringen par « faire-émerger », ce terme s'opposant d'emblée à l'émergence par le verbe « faire », qui implique un geste du sujet et ne dépend pas uniquement d'une qualité intrinsèque de ce qui émerge* » **Fuente especificada no válida..**

Dans les deux cas, l'écoute profonde explore, « se déplace » autour les objets sonores en explorant sa profondeur acoustique. Cela me ramène à l'idée que l'actant du dispositif audiovisuel est aussi un actant transformationnel. Si l'écoute est une activité, elle a également des étapes, elle se déroule comme un épisode. Ainsi, selon Krueger, la phase initiale de l'écoute profonde commence lorsque l'auditeur dirige l'attention à capturer la forme et la texture d'un trait sonore particulier. Donc, on rentre dans l'espace de la structure interne de la musique, dans ses dynamiques temporelles internes. L'auditeur peut commencer à écouter le contour de la mélodie, puis à suivre sa dynamique dans la mesure où elle dessine un « chemin narratif à travers le temps musical et l'espace tonal ». Dans l'écoute profonde, une mélodie est vécue comme quelque chose qui se déroule dans une dimension auditive spatialisée habitée par l'auditeur qui se déplace simultanément à travers d'elle; d'où la nature immersive de cette expérience.

La deuxième phase d'écoute profonde commence quand une sensibilité plus globale ou holistique est vécue. Une autre région de l'espace intérieur devient accessible: l'espace des relations musicales. En étant consciente des autres sons du morceau, l'appréciation de l'espace tonale est amplifiée permettant la perception de la texture, comprise comme la densité constituée par de multiples événements sonores simultanés. Par exemple, la mélodie et l'accompagnement non seulement se confondent en un seul continuum temporel, mais apparaissent comme occupants de différents emplacements spatiaux, en gardant à la fois leur individualité et leur relation mutuelle.

Dans la visu-audition de l'animation sonore l'écoute profonde arrive à cette phase lorsque le parcours par les traits sonores termine dans ce que Don Ihde appelle un *état de champ*, un espace totalisant dans lequel les différentes dimensions sonores sont intégrées dans un effet entourant comme ce décrit par Krueger pour la musique. Dans notre cas, ce champ d'état sonore fait que l'expérience visuelle soit également « élargie » dans la mesure où le flux d'images qui suivent le rythme et les progressions de son sont intégrés dans une expérience totalisante, qui ne cherche plus une similitude avec un monde (représenté), mais qu'*rentre dans un monde*. Il est facile d'associer cette description avec un état de transe. En effet, Judith Becker a associé l'écoute profonde à ce genre d'états de transe, et elle offre de nombreux exemples du rôle rituel de la musique dans les différentes cultures (BECKER, 2004).

Ainsi, l'actant de la visu-audition se transforme, il *change d'état* lors qu'il est en conjonction avec les mondes possibles de vibrations pures, où il devient lui-même pure vibration, chair résonnante.

### **Le flux de la vie et des logiques énergétiques**

Selon Mark Johnson, l'un des principaux représentants de la cognition incarnée, « La musique est significative parce qu'elle peut présenter le flux de l'expérience humaine, des émotions et de la pensée dans des formes concrètes et incarnées - et c'est le sens dans sa forme le plus profonde » (JOHNSON, 2008, p. 236)<sup>243</sup>.

Le dispositif animatique partage cette caractéristique, en proposant une exploration du son dont profite aussi de la musicalité (mais pas tout le temps). Dans mes analyses, j'ai essayé de ne pas faire des interprétations en termes de la théorie musicale, même lorsqu'il y avait des structures musicales, alors que j'ai cherché à me concentrer dans le niveau plus élémentaire du son lui-même, avant d'être structurée par le langage bien connu de la musique. En fait, je pense que la valeur de *Synchromie* ne se trouve pas dans sa composition musicale, mais dans la manipulation des caractéristiques structurelles du son lui-même, en tant que phénomène physique, et en correspondance avec les images proposées par McLaren.

Le son dessiné présente également « le flux de l'expérience humaine, des émotions et de la pensée », et il le fait d'une façon moins "chargée" par les formes culturelles musicales, en allant directement au corps comme une forme d'énergie qu'impacte et modifie une matière. Ainsi, s'il y a en effet un mode énergétique de se sentir et d'interpréter le cinéma, c'est cette expérience vivante de la synchrèse figurale où il n'y a aucune diégèse, mimesis ou histoire qui perturbe ou dirige la vibration du corps avec des images sonores et visuelles.

Dans son livre *Le son: Traité d'acoulogie*, Michel Chion propose de dépasser l'idée de «objet sonore» proposée par Schaeffer, car elle réduit les phénomènes sonores à l'étatisme des choses. En proposant une conception plus dynamique, proche de notre idée des sons conçus

---

<sup>243</sup> “Music is meaningful because it can present the flow of human experience, feeling, and thinking in concrete, embodied forms –and this is meaning in its deepest sense”.

comme des événements, Chion fait appelle à François Bayle. Ce compositeur propose aussi une conception dynamique du son:

« Qui dit dynamique implique la notion de force, d'énergie, et l'on peut écouter et manier les sons comme des phénomènes énergétiques et en mouvement, non comme des objets alignés et empilés sur les étagères du temps (...). C'est ainsi que Bayle a eu l'idée de considérer le son comme de l'énergie en mouvement et de classer les rapports d'enchaînement entre les sons non pas seulement comme des rapports abstraits de comparaison, de différence de degrés sur une échelle, mais aussi comme des rapports actifs de transmission, d'échange, de réaction ou de conflit d'énergie » (CHION, 2010, p. 4130 de 5607).

Bayle a identifié une série de « figures énergétiques », dont Chion ne cite qu'une: les "enchaînements en delta", du son "à l'envers", plus ou moins fréquents dans la musique concrète (tant Chion comme Bayle ont utilisé). Ce système présente un «cycle parfait» de « naissance-développement-résorption » (un schéma très proche de ce que j'ai identifié dans la structure aspectuelle des timbres de Mirai Mizue). Sur ces schèmes Chion dit:

Il semble pourtant que ces schèmes soient en nombre limité, parce qu'ils ne sont pas dessinés comme cela, comme des gribouillis capricieux tracés par la main de la nature ou celle d'un compositeur ou d'un musicien de jazz, mais plutôt en référence directe à des modèles naturels et physiologiques. C'est-à-dire qu'ils seraient la traduction sonore –ou plus, manifesterait l'existence sur un plan sonore- de processus plus généraux que notre expérience humaine nous a fait assimiler (CHION, 2010, p. 4147 de 5607).

En ce sens, on peut dire que le sens des formes audiovisuelles, et en particulier celles qui se manifestent sous la forme de la synchrèse figural que je viens de décrire ici, mais aussi dans les formes de personnages, des mondes et des récits figuratifs, sort du corps, passe par le corps et arrive au corps, comme le flux de la vie.

## Bibliographie

- Agamben, G. (2014). *Qu'est-ce qu'un dispositif ?* Paris: Rivages.
- Aloff, M. (2009). *Hippo in a Tutu: Dancing in Disney Animation*. New York: Disney Editions.
- Althusser, L. (1976). *Positions*. Paris: Les Éditions sociales.
- Angenot, V., Dondero, M. G., Joachim, G., & Shirkhodaei, S. (2013). Sémiotique de la communication en coprésence et à distance. Du textualisme à la sémiotique des pratiques. *Interfaces numériques. Volume 2 #3*, 531 - 567.
- Aumont, J. (1983). Le point de vue. *Communications #38*, 3-29.
- Aumont, J. (2011). *L'image*. Paris: Armand Colin.
- Barrès, P. (2006). *Le cinéma d'animation. Un cinéma d'expériences plastiques*. Paris: L'Harmattan.
- Bastiancich, A. (1997). *Norman McLaren. Précurseur des Nouvelles Images*. Paris: Dreamland éditeur.
- Baudry, J.-L. (1970). Effets idéologiques produits par l'appareil de base. *Cinéthique, # 7* - 8, 1 - 8.
- Baudry, J.-L. (1975). Le dispositif: approches métapsychologiques de l'impression de réalité. *Psychanalyse et cinéma. Communications, #23*, 56-72.
- Baudry, J.-L. (1978). *L'effet cinéma*. Paris: Éditions Albatros.
- Beauchamp, R. (2005). *Designing Sound for Animation*. Oxford: Focal Press.
- Becker, J. (2004). *Deep Listeners: Music, Emotion, and Trancing*. Bloomington: Indiana University Press.
- Beckman, K. (2014). Animating Film Theory: An Introduction. Dans K. (. BECKMAN, *Animating Film Theory* (pp. 1-36). Durham: Duke University Press.
- Bellour, R. (2009). *Le corps du cinéma. Hypnoses, émotions, animalités*. Paris: P.O.L.
- Bendazzi, G., Ceconello, M., & Michelone, G. (1995). *Coloriture : Voci, rumori, musiche nel cinema d'animazione*. Bologna: Pendragon .
- Bettetini, G. (1979). *Tempo del senso: la logica temporale dei testi audiovisivi*. Milano: Bompiani.
- Bettetini, G. (1996). *La conversación audiovisual. Problemas de la enunciación fílmica y televisiva*. Madrid: Cátedra.
- Bettetini, G. (2009). *Il timpano dell'occhio. Gli intrecci e i giochi dei suoni e delle immagini*. Milan: Studi Bompiani.

- Bionda, M. L. (1995). *Sound and Cartoons. Appunti per un approccio narratologico.*
- Blanco, D. (2003). *Semiótica del texto fílmico.* Lima: Universidad de Lima.
- Boillat, A. (2006). *La fiction au cinéma.* Paris: L'Harmattan.
- Bordron, J.-F. (2010). Quelques réflexions préliminaires en vue d'une sémiotique des sons. *Nouveaux Actes Sémiotiques*, <http://epublications.unilim.fr/revues/as/2832>.
- Bordwell, D. (1985). *Narration in the fiction film.* Wisconsin: The University of Wisconsin Press.
- Brémond, C. (1973). *Logique du récit.* Paris: Seuil.
- Brophy, P. (1991). The Animation of Sound. Dans A. Cholodenko, *The Illusion of Life. Essays on Animation* (pp. 67-12). Sidney: Power Publications/Australian Film Commission.
- Brunetta, G. P. (1970). *Forma e parola nel cinema.* Padua: Liviana.
- Buckland, W. (2004). *The Cognitive Semiotics of Film.* Cambridge: Cambridge University Press.
- Calveti, P. (1995). Look at me! I'm dancing! Danza e cartoon, il binomio indissolubile. Dans G. Bendazzi, M. Cecconello & G. Michelone, édés. *Coloriture. Voci, rumori, musiche nel cinema d'animazione.* Bologne: Pendragon, pp. 125-130.
- Camera, F. (1995). Una notte sul Monte Calvo. Un modello di trascrizione cinematografica. En G. Bendazzi, M. Cecconello, & G. Michelone, *Coloriture. Voci, rumori, musiche nel cinema d'animazione* (págs. 273 - 306). Bologne: Pendragon.
- Candia, R. (1995). Rumori di Carta : per una teoria del sonoro nell'animazione.
- Casetti, F. (1990). *D'un regard l'autre. Le film et son spectateur.* Lyon: Presses Universitaires de Lyon.
- Casetti, F. (2005). *Les Théories du cinéma depuis 1945.* Paris: Armand Colin.
- Châteauevert, J. (1996). *Des mots a l'image. La voix over au cinéma.* Paris/Québec: Méridiens Klincksieck/Nuit Blanche.
- Chion, M. (1983). *Guide des objets sonores. Pierre Schaeffer et la recherche musicale.* Paris: Buchet/Chastel, Institut National de l'Audiovisuel.
- Chion, M. (2003). *Un art sonore, le cinéma. Histoire, esthétique, poétique.* Paris: Cahiers du cinéma.
- Chion, M. (2008). *L'audio-vision. Son et image au cinéma.* Paris: Armand Colin.
- Chion, M. (2010). *Le son, traité d'acoulogie.* Paris: Armand Colin.
- Coyle, R. (2010). *Drawn to sound. Animation Film Music and Sonicity.* London: Equinox.



- Curtis, S. (1992). The sound of the early Warner Bros. cartoons. En R. Altman, *Sound Theory Sound Practice* (págs. 191 - 203). New York: Routledge.
- De Laurentis, T., & Heath, S. (1980). *The Cinematic Apparatus*. London: McMillan.
- De Nanteuil (ed.), P.-L. (2008). *Dictionnaire encyclopédique du son*. Paris: Dunod.
- Deleuze, G. (2003). *Deux régimes de fous*. Paris: Les éditions de minuit.
- Denis, S. (2011). *Le cinéma d'animation*. Paris: Armand Colin.
- Díaz, C. (2015). *Granos, píxeles y otras cosas del estilo. Aspectos plásticos de la fotografía cinematográfica*. Bogotá: Universidad Jorge Tadeo Lozano.
- Dobson, T. (2006). *The Film Work of Norman McLaren*. London: John Libbey.
- Donnelly, K. J. (2014). *Occult Aesthetics: Synchronization in Sound Film*. New York: Oxford University Press.
- Estay Stange, V. (2014). *Sens et musicalité : Les voix secrètes du symbolisme*. Paris: Editions Classiques Garnier.
- Floch, J.-M. (1983). Stratégies de communication syncrétique et procédures de syncrétisation. *Actes Sémiotiques*, 3 - 8.
- Floch, J.-M. (1985). *Petites mythologies de l'oeil et de l'esprit. Pour une sémiotique plastique*. Paris-Amsterdam: Hadès-Benjamins.
- Fontanille, J. (2003). *Sémiotique du discours*. Limoges: PULIM.
- Fontanille, J. (1989). *Les espaces subjectifs. Introduction à la sémiotique de l'observateur (discours - peinture - cinéma)*. Paris: Hachette.
- Fontanille, J. (1991). Aspectualisation, quantification et mise en discours. Dans J. Fontanille, *Le discours aspectualisé* (pp. 127-143). Limoges: PULIM/BENJAMINS.
- Fontanille, J. (1995). *Sémiotique du visible : des mondes de lumière*. Paris: PUF.
- Fontanille, J. (2008). *Pratiques sémiotiques*. Paris: PUF.
- Fontanille, J. (2010). Une sémiotique du son ? Remarques sur la constitution d'un plan d'immanence. *Nouveaux Actes Sémiotiques*, <http://epublications.unilim.fr/revues/as/2823>.
- Fontanille, J. (2011). *Corps et sens*. Paris: PUF.
- Fontanille, J. (2011). L'analyse du cours d'action : des pratiques et des corps. *SEMEN. Revue de sémio-linguistique des textes et discours*, 131-158.
- Fontanille, J. (2015). *Formes de vie*. Liège: Presses Universitaires de Liège.
- Fontanille, J., & Zilberberg, C. (2004). *Tensión y significación*. Lima: Fondo de desarrollo editorial - Universidad de Lima.

- Furniss, M. (2007). *Art in Motion. Animation Aesthetics*. Eastleigh: John Libbey Publishing.
- Gaudreault, A. (1999). *Du littéraire au filmique*. Paris: Nota Bene / Armand Colin.
- Gaudreault, A., & Jost, F. (1990). *Le récit cinématographique*. Paris: Éditions Nathan.
- Greimas, A. J. (1984). Semiotique figurative et semiotique plastique. *Actes Sémiotiques*, 5-24.
- Greimas, A. J., & Courtés, J. (1979). *Sémiotique. Dictionnaire raisonné de la théorie du langage*. Paris: Hachette.
- Greimas, A. J., & Courtés, J. (1993). *Sémiotique. Dictionnaire raisonné de la théorie du langage*. Paris: Hachette.
- Greimas, A. J., & Fontanille, J. (1991). Avant-Propos. Dans J. Fontanille, *Le discours aspectualisé* (pp. 5-16). Limoges: PULIM/BENJAMINS.
- Grodal, T. (2009). *Embodied Visions: Evolution, Emotion, Culture, and Film*. Oxford: Oxford University Press.
- Groupe  $\mu$ . (1992). *Traité du signe visuel : pour une rhétorique de l'image*. Paris: Seuil.
- Hard, B., Tversky, B., & Lang, D. S. (2006). Making Sense of Abstract Events. *Memory & Cognition*, 1221-1235.
- Hébert, L. (5 de octobre de 2016). *Signo. Site Internet des théories sémiotiques*. Petite sémiotique du rythme. Éléments de rythmologie: <http://www.signosemio.com/semiotique-du-rythme.asp>
- Huron, D. (2006). *Sweet Anticipation. Music and the Psychology of Expectation*. Cambridge: The MIT Press.
- Idhe, D. (2007). *Listening and Voice. Phenomenologies of Sound*. New York: State University of New York Press.
- Izvolov, N. (2014). From the History of Graphic Sound in the Soviet Union; or, Media without Medium. Dans L. Kaganovsky, & Salazkina, Masha, *Sound, Speech, Music in Soviet and Post-Soviet Cinema* (pp. 46 - 77). Bloomington: Indiana University Press.
- Jacquinet-Delaunay, G., & Monnoyer, L. (1999). Avant-propos. *Hermès*, 9 - 14.
- Johnson, M. (2008). *The Meaning of the Body*. Chicago: University of Chicago Press.
- Jost, F. (1987). *L'œil-camera: Entre film et roman*. Lyon: Presses universitaires de Lyon.
- Jost, F. (1992). *Un monde à notre image: Énonciation, cinéma, télévision*. Paris: Meridiens/Klincksieck.
- Jost, F. (2014). *Sous le cinéma, la communication*. Paris: VRIN.

- Jullier, L. (1995). *Les sons au cinéma et à la télévision. Précis d'analyse de la bande-son*. Paris: Armand Colin.
- Jullier, L. (2002). *Cinéma et cognition*. Paris: L'Harmattan.
- Jullier, L. (2012). *Analyser un film. De l'émotion à l'interprétation*. Paris: Flammarion.
- Kepley Jr., V. (1996). Whose Apparatus? Problems of Film Exhibition and History. Dans D. Bordwell, & N. Carroll, *Post-Theory: Reconstructing Film Studies* (pp. 533-549). Wisconsin: University of Wisconsin Press.
- Kessler, F. (2003). La cinématographie comme dispositif (du) spectaculaire. *Cinémas : revue d'études cinématographiques*, vol. 14, n° 1, 21 - 34.
- Krueger, J. (2011). Enacting Musical Content. Dans *Situated Aesthetics: Art beyond the Skin*. Exeter: Imprint Academic, pp. 63-85.
- Kurihara, U. (2005). Aesthetics of experimental animation. On dictatorial observation of Norman McLaren's *Synchromy* (1971). *International Journal of Asia Digital Art and Design*, 7 - 12.
- Lack, R. (1997). *Twenty Four Frames Under: A Buried History of Film Music*. London: Quartet Books.
- Langlois, P. (2012). *Les cloches d'Atlantis. Musique électroacoustique et cinéma. Archéologie et histoire d'un art sonore*. Paris: Éditions MF.
- Leman, M. (2008). *Embodied Music Cognition and Mediation Technology*. Cambridge: The MIT Press.
- Levin, D. Y. (2006). "Tones from out of Nowhere": Rudolph Pfenninger and the Archaeology of Synthetic Sound. Dans W. H. Kyong Chun, & T. Keenan, *New Media, Old Media. A history and Theory Reader* (pp. 33 - 79). New York: Routledge.
- Limoges, J.-P. (2012). La métalepse au cinéma. Aux frontières de la transgression. *Cinergie. Il cinema et le altre arti*, 126 - 144.
- Lucena Júnior, A. (2002). *Arte da animação. Técnica e estética através da historia*. Sao Paulo: Senac.
- Lytard, J.-F. (1994). *Des dispositifs pulsionels*. Paris: Galilée.
- Manovich, L. (2005). *El lenguaje de los nuevos medios de comunicación*. Barcelona: Paidós.
- Martin, A. (2000). On a Touché le Cinéma. Dans B. Clarens, *André Martin. Écrits sur l'animation. 1* (pp. 235 - 252). Paris: Dreamland éditeur.
- Martin, M. (1955). *Le langage cinématographique*. Paris: Les Éditions du CERF.
- McLaren, N. (1997). Note sur la musique Synthétique. Dans A. Bastiancich, *Norman McLaren: précurseur des Nouvelles Images* (pp. 158 - 162). Paris : Dreamland éditeur.

- McLaren, N. (1997). Quelques Notes sur le Son Animé Utilisé dans les Films : Now is the Time, A Phantasy, Voisins et Two Bagatelles. Dans A. Bastiancich, *Norman McLaren : Précurseur des Nouvelles Images* (págs. 163 - 168). Paris: Dreamland éditeur.
- McLaren, N. (1997). Son Animé sur Pellicule. Les Débuts. Dans A. Bastiancich, *Norman McLaren: Précurseur des Nouvelles Images* (pp. 154 - 157). Paris: Dreamland éditeur.
- Menarini, A. (1955). *Il cinema nella lingua, la lingua nel cinema*. Milan-Rome: Bocca.
- Meric, R. (2012). *Appréhender l'espace sonore. L'écoute entre perception et imagination*. Paris: L'Harmattan.
- Metz, C. (1964). "Le cinéma: langue ou langage?". *Communications*, #4, 52 - 90.
- Metz, C. (1971). *Langage et cinéma*. Paris: Larousse.
- Metz, C. (1975). Le film de fiction et son spectateur. *Psychanalyse et cinéma. Communications #23*, 108-135.
- Metz, C. (1977a). *Le signifiant imaginaire : psychanalyse et cinéma*. Paris: Inédit 1018.
- Metz, C. (1977b). *Essais Sémiotiques*. Paris : Klincksieck.
- Metz, C. (1991). *L'énonciation impersonnelle, ou le site du film*. Paris: Méridiens Klincksieck.
- Michelone, G. (1995). I love Singa. Per una fenomenologia del jazz-cartoon. Dans G. Bendazzi, M. Ceconello & G. Michelone, édés. *Coloriture. Voci, rumori, musiche nel cinema d'animazione*. Bologne: Pendragon, pp. 111-124.
- Mitry, J. (1963). *Esthétique et psychologie du cinéma. 1. Les structures*. Paris: Editions universitaires.
- Mitry, J. (1963). *Esthétique et psychologie du cinéma. 2. Les formes*. Paris: Editions universitaires.
- Mizue, M. (2010). Interview. (Inconnu, Intervieweur)
- Mizue, M. (2010, Août). Note sur Fantastic Cell. *Mirai Mizue Works 2003-2010 (Livret du DVD)*. Tokyo, Japon: CALF.
- Mollaghan, A. (2015). *The Visual Music Film*. Basingstoke: Palgrave Macmillan.
- Monet Claude. *La Rue Montorgueil*, 1878. Huile sur toile, 81 x 50 cm. Musée d'Orsay.
- Moritz, W. (2004). *Optical Poetry: The Life and Work of Oskar Fischinger*. Londres: John Libbey.
- Neupert, R. (2011). "We're Happy When We're Sad". Comedy, Gags, and 1930s Cartoon Narration. Dans D. Goldmark & C. Keil, edits. *Funny Pictures. Animation and Comedy in Studio-Era Hollywood*. Berkeley: University of California Press, pp. 93-108.

- Niño, D. (2015). *Elementos de semiótica agentiva*. Bogotá: Universidad Jorge Tadeo Lozano.
- O'Callaghan, C. (2007). *Sounds. A Philosophical Theory*. Oxford: Oxford University Press.
- Odin, R. (2000). *De la fiction*. Bruxelles: De Boeck Université.
- Patteson, T. (2015). *Instruments for New Music. Sound, Technology, and Modernism*. Oakland: University of California Press.
- Peeters, H., & Charlier, P. (1999). Contributions à une théorie du dispositif. *Hermès*, 15-23.
- Pelletier, J.-M. (29 de 07 de 2009). *The Birth of the Synthetic Voice*. Dans Jean-Marc Pelletier. Sound, art, media, education: <http://jmpelletier.com/tag/voice/>
- Pezzini, I. (2013). Stratégies de condensation dans les formes brèves. La forme brève comme genre médiatique, entre standardisation et créativité. Dans S. Péreineau, *Les formes brèves audiovisuelles. Des interludes aux productions web* (pp. 29-46). Paris: C.N.R.S. Editions.
- Radvansky, G. A., & Zacks, J. M. (2014). *Event Cognition*. Oxford: Oxford University Press.
- Rae Hark, I. (. (2001). *Exhibition, The Film Reader*. New York: Routledge.
- Rondolino, G. (1997). Introduction. En A. Bastiancich, *Norman McLaren. Précurseur des Nouvelles Images* (págs. 13 - 26). Paris: Dreamland Éditeur.
- Schaeffer, P. (2003 [1966]). *Tratado de los objetos musicales*. Madrid: Alianza Editorial.
- Selvini, M. (1995). Caro Amico! I suoi programmi... Riflessioni su musica classica e cinema d'animazione. Dans G. Bendazzi, M. Cecconello & G. Michelone, eds. *Coloriture. Voci, rumori, musiche nel cinema d'animazione*. Bologne: Pendragon, pp. 101-110.
- Sirois-Trahan, J.-P. (2003). Dispositif(s) et réception. *Cinémas : revue d'études cinématographiques*, vol. 14, n° 1, 149 - 176.
- Stam, R. (2000). *Film Theory: an Introduction*. Massachusetts: Wiley-Blackwell.
- Stam, R., Burgoyne, R., & Flitterman-Lewis, S. (1999). *New Vocabularies in Film Semiotics: Structuralism, Post-Structuralism and Beyond*. London-New York: Routledge.
- Tarasti, E. (1994). *A Theory of Musical Semiotics*. Indianapolis: Indiana University Press.
- Varela, F., Thompson, E. & ROSCH, E. (1993). *L'inscription corporelle de l'esprit. Sciences cognitives et expérience humaine*. Paris: Le Seuil.

- Virilio, P. (1984). *Guerre et cinéma I - Logistique de la perception*. Paris: Cahiers du Cinéma / Éditions de l'Étoile.
- Waller, G. A. (2001). *Moviegoing in America: A Sourcebook in the History of Film Exhibition*. Malden: Wiley-Blackwell.
- Wells, P. (2010). *Understanding Animation*. New York: Routledge.
- Wells, P. (2011). The Chaplin Effect. Ghosts in the Machine and Animated Gags. Dans D. Goldmark & C. Keil, edits. *Funny Pictures. Animation and Comedy in Studio-Era Hollywood*. Berkeley: University of California Press, pp. 15-28.
- Zacks, J. M. (2015). *Flicker. Your Brain on Movies*. New York: Oxford University Press.
- Zielinski, S. (1999). *Audiovisions. Cinema and Television as Entr'Actes in History*. Amsterdam: Amsterdam University Press.

# Filmographie

1. Silly Symphony: The Skeleton Dance (Walt Disney, 1929)

La Danse macabre

DVD: Silly symphonies - les contes musicaux

Disney, 2015

<https://www.youtube.com/watch?v=vOGhAV-84iI>

2. Steamboat Willie (Walt Disney, 1928)

DVD: Mickey Mouse in Black and White - The Classic Collection

Mickey mouse - les années noir et blanc, vol. 1 : 1928 à 1935

Walt Disney Studios Home Entertainment, 2002

<https://www.youtube.com/watch?v=IFGqZUTqorU>

3. Une nuit sur le Monte Chauve (Alexandre Alexeïeff, 1933)

DVD: Alexeïeff. Le cinéma épinglé

Cinédoc, 2006

<https://www.youtube.com/watch?v=wYbjW7XrWDo>

4. Who Framed Roger Rabbit (Robert Zemeckis, 1988)

Qui veut la peau de Roger Rabitt

DVD: Touchstone Home Video, 2003

5. What Price Fleadow (Tex Avery, 1948)

Homer la Puce

DVD: Tex Avery, volume 2

Warner Home Video, 2004

<https://vimeo.com/146027299>

6. Gerald McBoing Boing (UPA-Robert Cannon, 1950)

DVD: Cartoon Adventures Starring Gerald McBoing Boing

Sony Pictures Home Entertainment, 2006

<https://www.youtube.com/watch?v=BI12ZsTR3Mo>



7. Zoom and Bored (Chuck Jones, 1957)

DVD: LOONEY TUNES GOLDEN COLLECTION - VOLUME 2 (Disque 2)

Warner Home Video, 2005

<http://www.b99.tv/video/zoom-bored/>

8. Balance (Christoph et Wolfgang Lauenstein, 1989)

DVD: Zeitgenossen – Deutscher Animationsfilm der Gegenwart (German Animation Collection)

Absolut MEDIEN, 2006

<https://www.youtube.com/watch?v=7wJj58aLvdQ>

9. La joie de vivre, (Anthony Gross et Hector Hoppin, 1934)

DVD: Du praxinoscope au cellulo. Un demi-siècle de cinéma d'animation en France (1892-1948).

CNC, 2007

<http://www.youtube.com/watch?v=Y2gxu7mD-5c>

10. Neighbors (Norman McLaren, 1952)

Voisins

DVD: Norman McLaren. L'intégrale (Disque 5)

ONF / Image Entertainment, 2006

[https://www.youtube.com/watch?v=i\\_2\\_VYY14-Y](https://www.youtube.com/watch?v=i_2_VYY14-Y)

11. Fantasia (Walt Disney, 1940)

DVD: Walt Disney France, 2010

<https://www.youtube.com/watch?v=r7gLLIv4ito>

12. Motion Painting No. 1 (Oskar Fischinger, 1947)

DVD: Oskar Fischinger: Ten Films

Center for Visual Music, 2006

[https://vk.com/video2804234\\_168196055](https://vk.com/video2804234_168196055)

13. La Vieille Dame et les pigeons (Sylvain Chomet, 1998)

DVD: La Vieille dame et les pigeons et 3 autres bijoux d'humour noir  
France Television Distribution. Paris, 2001  
<https://www.youtube.com/watch?v=MqhVcRrrauY>

14. Les Triplettes de Belleville (Sylvain Chomet, 2003)  
DVD: France Télévisions Distribution, Paris, 2004

15. The Brave Little Toaster (Jerry Rees, 1987)  
Le Petit Grille-pain courageux  
DVD: Seven7 Editions, 2004

16. Bugs Bunny Rides Again (Fritz Freleng, 1948)  
DVD: Looney Tunes: Golden Collection Vol. 2 (Disque 1)  
Warner Home Video  
<https://vimeo.com/54515426>

17. Jasper in a Jam (George Pal/Duke Goldstone, 1946)  
Jasper pris au piège  
DVD: Puppetoon. Le film.  
Pathé! 2005  
<https://www.youtube.com/watch?v=TYF7Zx8G5lg>

18. La belle et la bête (Kirk Wise et Gary Trousdale, 1991)  
DVD: Walt Disney France, 2015

19. Peter Pan (Clyde Geronimi, Wilfred Jackson, Hamilton Luske, 1953)  
DVD: Walt Disney France, 2012  
Séquence - The Flight to Neverland :  
<https://www.youtube.com/watch?v=zf00mEe9EOs>

20. Three little pigs (Burt Gillett, 1933)  
Les Trois Petits Cochons  
DVD: Silly symphonies - les contes musicaux  
Disney, 2015

21. Playful Pluto (Burt Gillett, 1934)

Pluto jongleur

DVD: Intégrale pluto. Les années 1931 À 1947

Disney, 2015

<https://www.youtube.com/watch?v=cmbDTG7lMzs>

22. The Band Concert (Wilfred Jackson, 1935)

DVD: Mickey Mouse, les années couleur - 1ère partie : les années 1935 à 1938

Walt Disney France, 2011

23. Rhapsody Rabbit (Friz Freleng, 1946)

DVD: Looney Tunes: Golden Collection Vol. 2 (Disque 4)

Warner Home Video, 2004

<http://www.dailymotion.com/video/x2n7vd9>

24. Red Hot Riding Hood (Tex Avery, 1943)

DVD: LooneyTunes Platinum Collection - Volume 2 (Disque 3)

Warner Home Video, 2012

[http://www.dailymotion.com/video/x262srt\\_tex-avery-mgm-1943-05-08-red-hot-riding-hood\\_fun](http://www.dailymotion.com/video/x262srt_tex-avery-mgm-1943-05-08-red-hot-riding-hood_fun)

25. Screwball Squirrel (Tex Avery, 1944)

DVD: LooneyTunes Platinum Collection - Volume 2 (Disque 3)

Warner Home Video, 2012

<https://vimeo.com/145808279>

26. Flip the Frog - The Cuckoo Murder Case (Ub Iwerks, 1930)

DVD: Coffret Cartoons, vol.1 : Flip The Frog

MK2, 2004

<https://www.youtube.com/watch?v=NEjR3jrW-mA>

27. What's Opera, Doc ? (Chuck Jones, 1957)

DVD: Looney Tunes: Golden Collection Vol. 2

Warner Home Video, 2005

<https://vimeo.com/115773980>

28. The Herring Murder Mystery (Dun Roman, 1943, 1944?)

<https://www.youtube.com/watch?v=q3uOapx20hY>

29. Santa's Workshop (L'atelier du père Noël) (Wilfred Jackson, 1932)

L'Atelier du Père Noël

<https://www.youtube.com/watch?v=6u-2PjzUXQI>

30. Tönendes ABC (László Moholy-Nagy, 1932/1933)

Son ABC

Scanné par Jan Thoben

<https://vimeo.com/114001563>

31. Vor (Nikolai Voinov, 1934)

Voleur

<https://www.youtube.com/watch?v=Z7Zb4rso82M>

32. Tönende Handschrift I, II et III (Rudolf Pfenninger, 1932)

DVD: Animierte Avantgarde – Der künstlerische Animationsfilm im Deutschland der 20er und 30er Jahre

Absolut MEDIEN, 2011

33. Synchronie (Norman McLaren, 1971).

DVD: Norman McLaren. L'intégrale (Disque 5)

ONF / Image Entertainment, 2006

<https://www.youtube.com/watch?v=UmSzc8mBJCM>

34. Dots (Norman McLaren, 1940).

DVD: Norman McLaren. L'intégrale (Disque 2)

ONF / Image Entertainment, 2006

<https://www.youtube.com/watch?v=E3-vsKwQOCg>

35. Loops (Norman McLaren, 1940).

DVD: Norman McLaren. L'intégrale (Disque 2)

ONF / Image Entertainment, 2006

<https://www.youtube.com/watch?v=6JvOqeEtFRY>

36. À la Pointe de la Plume (Don Peters et Lorne Batchelor, 1951)

DVD: Norman McLaren. L'intégrale (Disque 5)

ONF / Image Entertainment, 2006

37. Now is the Time (Norman McLaren, 1951)

BLURAY 3D: 3D Rarities

3D Film Archive/Flicker Alley, 2015

38. A Phantasy (Norman McLaren, 1952)

DVD: Norman McLaren. L'intégrale (Disque 2)

ONF / Image Entertainment, 2006

[https://www.youtube.com/watch?v=fjl0i\\_p\\_pow](https://www.youtube.com/watch?v=fjl0i_p_pow)

39. Two Bagatelles (Norman McLaren, 1952)

DVD: Norman McLaren. L'intégrale (Disque 3)

ONF / Image Entertainment, 2006

<https://www.youtube.com/watch?v=IcBOuA9GbxE>

40. Experimental film (Arthur Lipsett, 1962)

DVD: Norman McLaren. L'intégrale (Disque 6)

ONF / Image Entertainment, 2006

41. L'animateur musicien (Éric Barbeau, 2004)

DVD: Norman McLaren. L'intégrale (Disque 5)

ONF / Image Entertainment, 2006

42. The eyes Hears, the Ears Sees (Gavin Millard, 1970)

DVD: Norman McLaren. L'intégrale (Disque 6)

ONF / Image Entertainment, 2006

43. McLaren on McLaren (Grant Munro, 1983)

DVD: Norman McLaren. L'intégrale (Disque 6)

ONF / Image Entertainment, 2006

44. Rythmic, (Norman McLaren, 1956)

DVD: Norman McLaren. L'intégrale (Disque 6)

ONF / Image Entertainment, 2006

<https://www.youtube.com/watch?v=Uj8ETfWyE1s>

45. Mosaïque (Norman McLaren, 1965)

DVD: Norman McLaren. L'intégrale (Disque 2)

ONF / Image Entertainment, 2006

[https://www.youtube.com/watch?v=IwA\\_L9otKm8](https://www.youtube.com/watch?v=IwA_L9otKm8)

46. Test A pour Synchronie (Norman McLaren, fin des années soixante)

DVD: Norman McLaren. L'intégrale (Disque 5)

ONF / Image Entertainment, 2006

47. Test B pour Synchronie (Norman McLaren, fin des années soixante)

DVD: Norman McLaren. L'intégrale (Disque 5)

ONF / Image Entertainment, 2006

48. Fantastic Cell (Mirai Mizue, 2003)

DVD: Mirai Mizue. Works 2003-2010. Japanese Independent Animators Vol. 01  
CALF, 2010

49. Interview (Mirai Mizue, 2010).

DVD: Mirai Mizue. Works 2003-2010. Japanese Independent Animators Vol. 01  
CALF, 2010

50. Trip!-Trap!, (Mirai Mizue, 2005)

DVD: Mirai Mizue. Works 2003-2010. Japanese Independent Animators Vol. 01  
CALF, 2010

51. Lost Utopia, (Mirai Mizue, 2007)

DVD: Mirai Mizue. Works 2003-2010. Japanese Independent Animators Vol. 01  
CALF, 2010

52. Jam, (Mirai Mizue, 2009)  
DVD: Mirai Mizue. Works 2003-2010. Japanese Independent Animators Vol. 01  
CALF, 2010
53. Tatamp, (Mirai Mizue, 2011)  
DVD/BLURAY: L'Animation Independante Japonaise, Volume 1  
Les films du paradoxe/Carte Blanche/CNC, 2013
54. And And, (Mirai Mizue, 2011)  
DVD/BLURAY: L'Animation Independante Japonaise, Volume 2  
Les films du paradoxe/Carte Blanche/CNC, 2014
55. Wonder, (Mirai Mizue, 2012)  
DVD/BLURAY: L'Animation Independante Japonaise, Volume 2  
Les films du paradoxe/Carte Blanche/CNC, 2014
56. The Baby Birds Of Norman McLaren (Mirai Mizue, 2014)  
<https://vimeo.com/85837196>
57. Timbre A-Z (Mirai Mizue, 2011)  
<https://vimeo.com/31332018>
58. A long day of timbre (Mirai Mizue, 2011)  
<https://vimeo.com/20262614>
59. Animation-making process (Mirai Mizue, 2011)  
<https://vimeo.com/20222436>
60. Animation-making process - Complete Version (Mirai Mizue, 2011)  
<https://vimeo.com/31382034>



# Annexes

## Annexe 1

Description détaillée des transformations et changements sonores et visuels de *Synchromie* (Norman McLaren, 1971).

Dispositif matériel de réception édicté pour l'analyse :

Le film a été audio visionné pendant plusieurs séances. La plupart de ces séances furent à la nuit, sans lumière dans la salle. Les dernières séances ont été avec de la lumière artificielle. Le film a été audio visionné sur ordinateur.

Ordinateur : Toshiba Satellite P55t-A 5105SL

Écran :

TruBrite® TFT LCD. 15.6 pouces.

Aspect Ratio : 16 – 9

Résolution d'origine : 1920 x 1080 (FHD)

Carte son : DTS Studio Sound™

Casques : Marshall MAJOR

Transducteur : 40mm, dynamique

Impédance : 64 ohm

Plage de fréquence : 10 Hz – 20 kHz

Sensibilité : 100 mV @ 1kHz = 99 dB SPL

Film

Norman McLaren – L'intégrale (6 DVDs).

Une production de l'Office national du film de Canada – 2006

DVD 5

Format originel : 35 mm<sup>244</sup>

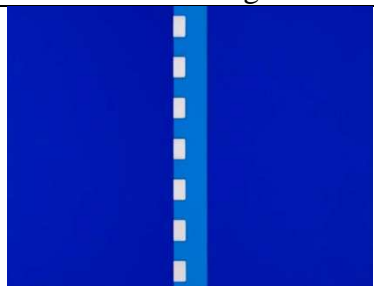
Son : Original (mono). Dolby digital 5.1. Dolby Surround<sup>245</sup>.

Durée : 7 min 26 s

Note technique :

Cartes de sons synthétiques photographiés sur la pellicule.

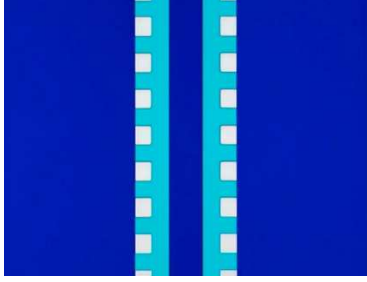
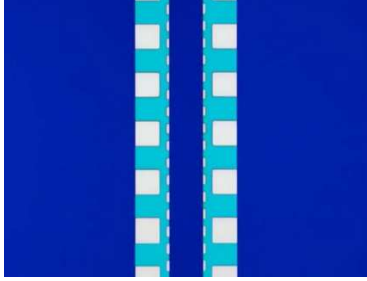
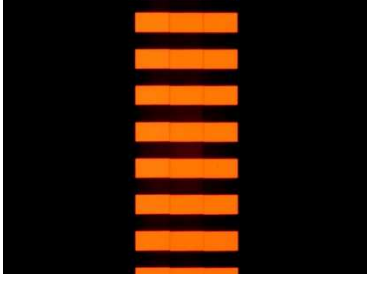
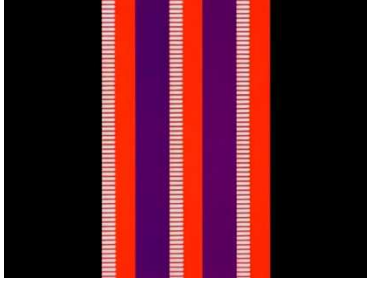
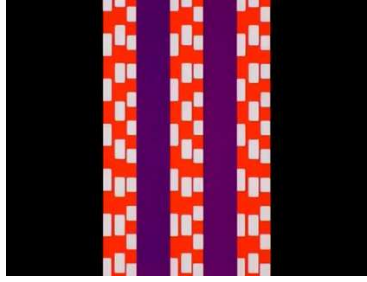
Mêmes cartes de sons photographiées sur la piste sonore de la pellicule

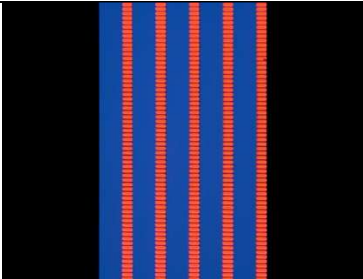
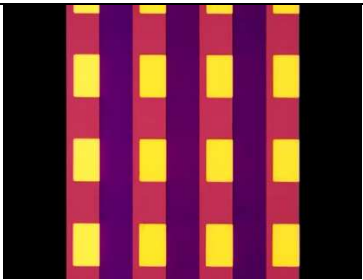
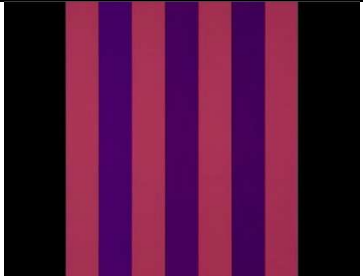
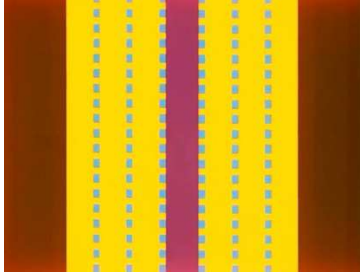

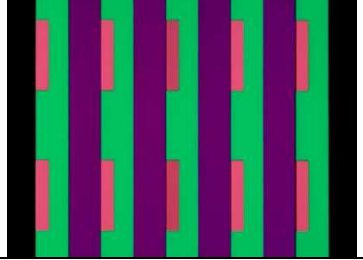
Bande image	Temps – Description	Bande-son
	0:21  Debout du film. On voit une seule colonne, représentation de la carte sonore.	Une ligne sonore.  Effet d'écho et de « profondeur sonore ». Augmentation, diminution et augmentation du volume (effet de distance et proximité).

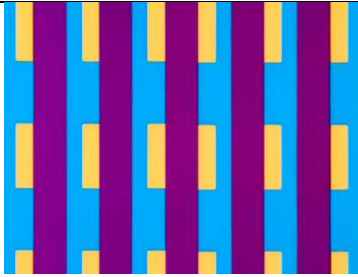
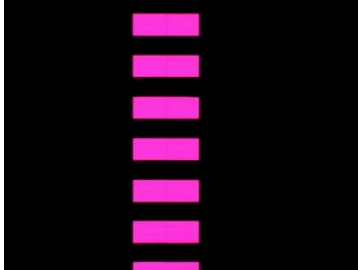
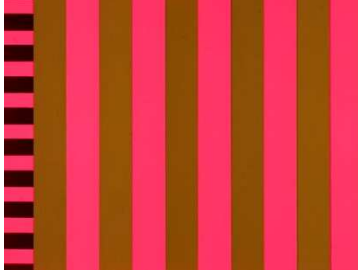
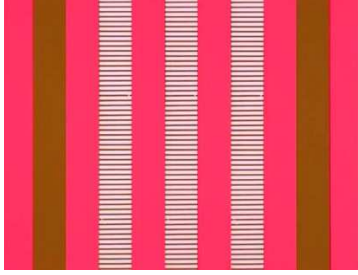
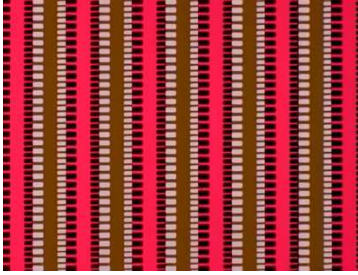
---

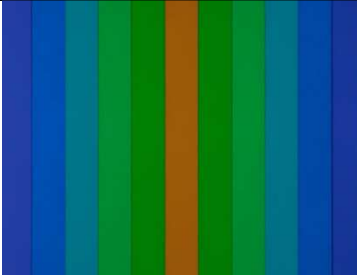
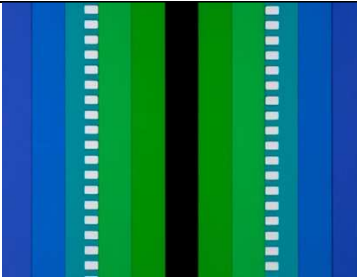
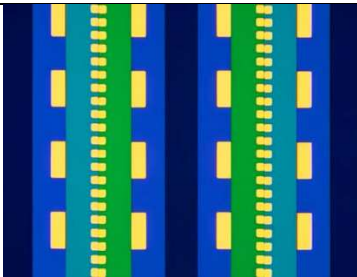
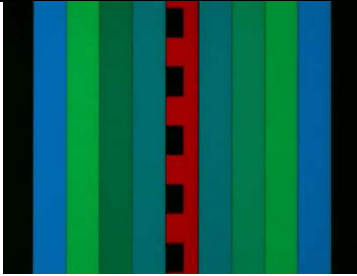
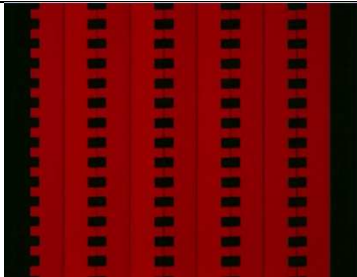
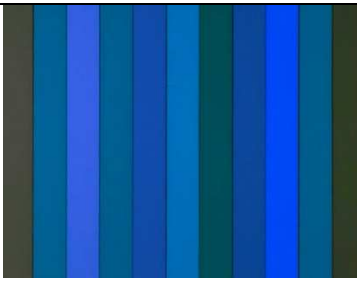
<sup>244</sup> Voici un fragment du commentaire technique du livret de la collection DVD sur la restauration des films pour l'édition intégrale : « La décision de placer l'image à l'intérieur d'un cadre noir fut prise afin de mettre en valeur les portions d'animation situées aux limites d'un cadre 35 mm. En effet, tout au long de sa carrière, McLaren a travaillé sur la totalité de la surface de la pellicule, plaçant fréquemment des éléments graphiques signifiants aux limites du cadre. Ces éléments étaient exclus lors des précédents transferts vidéo, le format 4 :3 des écrans traditionnels de télévision ne permettant pas le respect de l'intégralité du travail. En rapetissant légèrement l'image et en la plaçant à l'intérieur d'un cadre noir, nous donnons la possibilité aux spectateurs de poser un regard nouveau sur les films de Norman McLaren, plus conforme à la volonté originelle du cinéaste ».

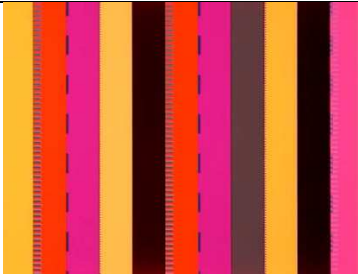
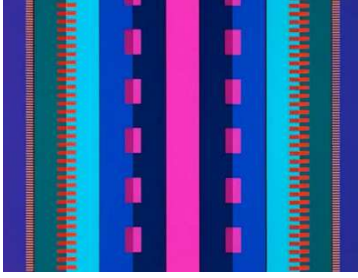
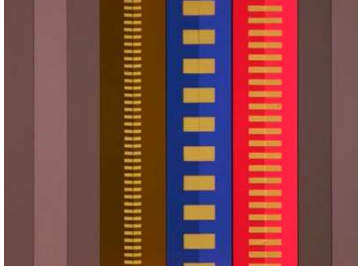
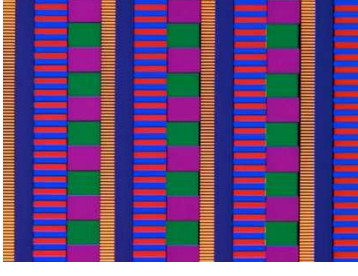
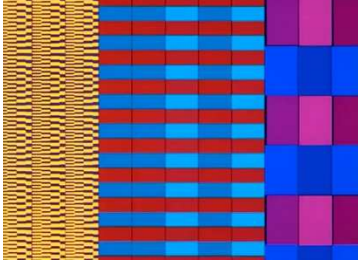
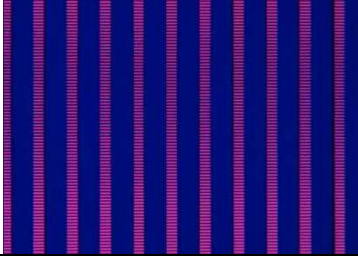
<sup>245</sup> Le film a été audio visionné à plusieurs reprises avec le son original monophonique. J'ai audio visionné le film aussi avec le son restauré et reconstruit par L'Office national du film de Canada avec les systèmes dolby digital et dolby 5.1., sur une *Home-Theater* Sony, avec son 5.1. Néanmoins, pour l'analyse, j'ai opté par le son original sur ordinateur avec du casque. Les effets spatiaux décrits provient de ce format. Voici une citation du livret de la collection DVD, sur la restauration sonore : « un travail de restauration sonore a été effectué sur tous les films sonores de McLaren, qui ont ensuite été remixés en mono, en Dolby 2.0 ainsi qu'en Dolby 5.1. En incluant les pistes mono dans la présente édition, nous avons voulu respecter l'intégrité des bandes originales. Donald McWilliams a supervisé ces opérations ».

	<p>1:20</p> <p>Deux colonnes apparaissent.</p>	<p>Une ligne sonore.</p> <p>Augmentations et diminutions du volume alternées (effet de distance et proximité).</p>
	<p>1:31</p> <p>Dans chaque colonne il y a deux files de carrés.</p>	<p>Deux lignes sonores (la plus forte à une vitesse plus élevée).</p> <p>Augmentations et diminutions du volume alternées (effet de distance et proximité).</p>
	<p>2:04</p> <p>Chaque occurrence de la note, apparaît une barre plus épaisse (composé de trois barres attachées) d'une seule couleur sur fond noir.</p>	<p>Une ligne sonore. Notes espacées apparaissent (aiguës et graves alternées).</p>
	<p>2:13</p> <p>Multiplication des colonnes. Trois (rouges) avec interstices (pourpres).</p>	<p>Effet d'augmentation. Accélération. Notes aiguës.</p>
	<p>2:56</p> <p>Accord visuel entre les trois colonnes. Chaque colonne avec trois lignes de carrés blancs.</p>	<p>Accords sonores. Trois accords trois fois.</p>

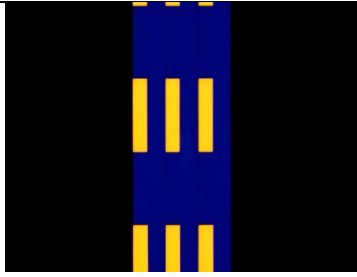
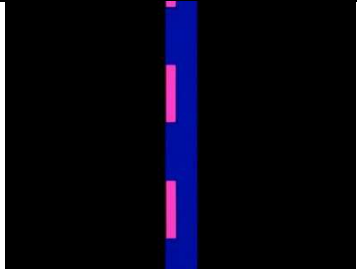

	<p>3:03</p> <p>Toutes les cinq colonnes du même couleur : bleus avec carrés orange.</p>	<p>Sons aigus avec vibration soutenue décroissante. Réverbération qui donne l'effet de plusieurs pistes, mais il n'y a que une.</p> <p>Disparition de l'écho pendant les aigus soutenus.</p>
	<p>3:08</p> <p>Multiplication des colonnes (sept). Accélération du mouvement.</p>	<p>Accélération sonore.</p>
	<p>3:13</p> <p>Colonnes sans carrés.</p>	<p>Silence.</p>
	<p>3:35</p> <p>Les sept colonnes sont organisées en deux méga-colonnes jeunes.</p>	<p>Deux lignes sonores (aiguës).</p> <p>Disparition de l'écho pendant les aigus soutenus.</p>
	<p>3:40</p> <p>Multiplication des colonnes (neuf).</p>	
	<p>3:41</p>	<p>Une seule ligne sonore.</p>

	<p>3:54</p> <p>Multiplication des colonnes (onze).</p> <p>4:01. Écran noir.</p>	
	<p>4:01</p> <p>Deux colonnes de la même couleur sur fonde noire. Intermittence et changement de couleur.</p>	<p>Notes aiguës et graves alternées.</p>
	<p>4:10</p> <p>Des carrés noirs sur les colonnes rossées apparaissent et se déplacent de gauche à droite. Sur les colonnes ocres apparaissent des carrés blancs et se déplacent de droite à gauche.</p>	<p>Sons aigus synchronisés avec les carrés noirs apparaissent en intermittence. Sons graves synchronisés avec les carrés blancs apparaissent en intermittence.</p> <p>Disparition de l'écho.</p>
	<p>4:24</p> <p>Mouvement des carrés blancs du centre à la périphérie. Alternance : colonne centrale, colonnes latérales (en symétrie).</p>	
	<p>4:32</p> <p>Deux lignes de carrés sur les colonnes.</p>	<p>Double ligne sonore.</p>

	<p>4:37 Silence</p>	<p>Une seule ligne sonore. Alternance note / silence.</p>
	<p>4:42</p> <p>Carrés blancs se déplacent des extrêmes au centre.</p>	<p>Une seule ligne sonore alternée avec des silences.</p>
	<p>4:51</p> <p>Les colonnes s'organisent en deux méga-colonnes.</p>	<p>Double ligne sonore. 5:05. La ligne sonore grave décroît jusqu'à disparaître. 5:12. La deuxième ligne apparaît une fois plus.</p>
	<p>5:18</p>	<p>Alternance de sons graves et aigus. Deux lignes sonores.</p>
	<p>5:24</p> <p>Les colonnes deviennent rouge obscur. Mouvement « centrifuge » du rouge. Cadres noirs en intermittence.</p>	
	<p>5:30.</p> <p>Changement des couleurs (de chauds aux froids)</p>	<p>Changement de hauteurs (vers les aigus).</p> <p>Alternance de sons graves et aigus (deux lignes sonores).</p> <p>Retour de l'écho.</p>

	<p>5:53</p> <p>Changement des couleurs (de froids aux chauds)</p> <p>Alternance des chauds et des froids</p>	Accélération du rythme
	<p>5:59</p> <p>Apparition de plusieurs carrés au-dessus des barres</p>	
	<p>6:15</p> <p>Trois colonnes saturées au-dessus des colonnes insaturées. Déplacement des colonnes saturées du centre à la périphérie.</p>	Trois lignes sonores apparaissent.
	<p>6:49</p> <p>Multiplication des carrés au-dessus des colonnes. Combinaison polychrome.</p>	Les trois lignes sonores ont des rythmes différents.
	<p>7:07</p> <p>Les colonnes sont organisées en trois groupes.</p>	Accélération musicale.
	<p>7:14</p> <p>Réduction progressive des colonnes (toutes avec fond bleu)</p>	Réduction progressive des lignes sonores, de trois à une.



	<p>7:18</p> <p>Trois colonnes</p>	<p>Alternance (très espacée) entre son et silence</p>
	<p>7:20</p> <p>Une colonne (éphémère)</p> <p>Fond noir</p>	<p>Silence court (moins d'une seconde)</p>
	<p>7:21</p> <p>Génériques</p>	<p>Accord multiple des sons, décroissance progressive du volume, réverbération, extinction</p>

## Annexe 2

Relations de Synchronisme (ODIN, 1990), succession, alternance et chevauchement dans *Timbre A to Z* (Mirai Mizue, 2011).

Dispositif matériel de réception énoncé pour l'analyse :

Le film a été audio visionné pendant plusieurs séances. La plupart de ces séances furent à la nuit, sans lumière dans la salle. Les dernières séances ont été avec de la lumière artificielle. Le film a été audio visionné sur ordinateur.

Ordinateur : Toshiba Satellite P55t-A 5105SL

Écran :

TruBrite® TFT LCD. 15.6 pouces.

Aspect Ratio : 16 – 9

Résolution d'origine : 1920 x 1080 (FHD)

Carte son : DTS Studio Sound™

Casques : Marshall MAJOR

Transducteur : 40mm, dynamique

Impédance : 64 ohm

Plage de fréquence : 10 Hz – 20 kHz

Sensibilité : 100 mV @ 1kHz = 99 dB SPL

Film

Le film a été audiovisionné directement sur le site internet Vimeo de l'artiste. Il n'y pas d'spécifications techniques (<https://vimeo.com/31332018>).

Durée : 13 min

Durée de chaque épisode : 30 s

Aspect Ratio : 16 – 9

Son : Stéréo

<b>Timbre A</b>	
Désynchronisme	Entre A <sub>1</sub> et A <sub>1</sub>
<b>Timbre B</b>	
Synchronisme	Entre B <sub>1</sub> et B <sub>3</sub>
A-rythmie	Entre B <sub>1</sub> et B <sub>2</sub>
<b>Timbre C</b>	
Désynchronisme	Entre C <sub>1</sub> et C <sub>2</sub>
Non-synchronisme	Entre C <sub>1</sub> (ralentie) et C <sub>2</sub>
<b>Timbre D</b>	
Non-synchronisme	Entre D <sub>1</sub> et D <sub>2</sub>
<b>Timbre E</b>	
Succession	Entre E <sub>1</sub> (4) et E <sub>2</sub> (1) Entre E <sub>3</sub> (4) et E <sub>4</sub> (1) Entre E <sub>5</sub> (4) et E <sub>6</sub> (1) Entre E <sub>2</sub> (4) et E <sub>3</sub> (1)
Chevauchement	Entre E <sub>4</sub> (4), E <sub>5</sub> (1), E <sub>1</sub> (1), E <sub>2</sub> (1), E <sub>6</sub> (1), E <sub>3</sub> (1), E <sub>4</sub> (1), E <sub>5</sub> (1), E <sub>1</sub> (1), E <sub>2</sub> (1), E <sub>6</sub> (1), E <sub>3</sub> (1)-E <sub>4</sub> (1)E <sub>5</sub> (1), E <sub>1</sub> , E <sub>3</sub> , E <sub>4</sub> , E <sub>5</sub> , E <sub>1</sub> , E <sub>2</sub> , E <sub>6</sub>
Simultanéité	E <sub>1</sub> , E <sub>2</sub> , E <sub>3</sub> , E <sub>4</sub> , E <sub>5</sub> , E <sub>6</sub>
<b>Timbre F</b>	
Synchronisme	F <sub>1</sub> , F <sub>2</sub> et F <sub>3</sub>
<b>Timbre G</b>	
Succession	G <sub>1</sub> et G <sub>2</sub>
Désynchronisme	G <sub>1</sub> et G <sub>2</sub>
<b>Timbre H</b>	
Chevauchement	H <sub>1</sub> et H <sub>1</sub>
<b>Timbre I</b>	
Désynchronisme	I <sub>1</sub> x 6

Timbre J	
Succession	J <sub>1</sub> et J <sub>2</sub>
Non-synchronisme	J <sub>1</sub> et J <sub>2</sub>
A-rythmie	J <sub>1</sub> et J <sub>2</sub>
Timbre K	
Synchronisme	K <sub>1</sub> et F <sub>3</sub>
Timbre L	
Désynchronisme	L <sub>1</sub> et L <sub>2</sub> L <sub>3</sub> et L <sub>2</sub> L <sub>1</sub> et L <sub>2</sub> L <sub>3</sub> et L <sub>2</sub> L <sub>1</sub> L <sub>2</sub> et L <sub>3</sub>
Timbre M	
Non-synchronisme A-rythmie	M <sub>1</sub> , C <sub>1</sub> et C <sub>2</sub>
Timbre N	
Succession	N <sub>1</sub> x 5 et B <sub>1</sub> x 1 N <sub>1</sub> x 5 et D <sub>1</sub> x 1 N <sub>1</sub> x 5 et B <sub>1</sub> x 1
Désynchronisme	N <sub>1</sub> et B <sub>1</sub> N <sub>1</sub> et B <sub>1</sub> + D <sub>1</sub>
Timbre O	
Synchronisme	O <sub>1</sub> et O <sub>2</sub> O <sub>1</sub> , O <sub>2</sub> et G <sub>1</sub> O <sub>1</sub> , O <sub>2</sub> , G <sub>1</sub> et G <sub>2</sub>
Timbre P	
Succession	P <sub>1</sub> et P <sub>2</sub> J <sub>2</sub> et P <sub>2</sub> F <sub>2</sub> et P <sub>2</sub>
Chevauchement	F <sub>3</sub> et P <sub>2</sub>
Timbre Q	
Synchronisme (?)	Q <sub>1</sub> , B <sub>1</sub> , B <sub>2</sub> , G <sub>1</sub> , L <sub>1</sub> et J <sub>1</sub>

Désynchronisme (?)	Q <sub>1</sub> , B <sub>1</sub> , B <sub>2</sub> , G <sub>1</sub> , L <sub>1</sub> et J <sub>1</sub>
<b>Timbre R</b>	
Synchronisme	R <sub>1</sub> et R <sub>2</sub>
Désynchronisme	R <sub>1</sub> +R <sub>2</sub> et R <sub>3</sub>
<b>Timbre S</b>	
A-rythmie	B <sub>2</sub> et A <sub>1</sub> B <sub>2</sub> et F <sub>3</sub> B <sub>2</sub> , A <sub>1</sub> et F <sub>3</sub>
<b>Timbre T</b>	
Succession	T <sub>1</sub>
Chevauchement	T <sub>1</sub>
<b>Timbre U</b>	
Succession	U <sub>1</sub> et C <sub>1</sub>
<b>Timbre V</b>	
Succession	V <sub>1</sub> et Q <sub>1</sub>
Synchronisme	K <sub>1</sub> et O <sub>2</sub> K <sub>1</sub> , L <sub>1</sub> , Q <sub>1</sub> et O <sub>2</sub>
A-rythmie	V <sub>1</sub> et K <sub>1</sub>
Non-synchronisme + A-rythmie	K <sub>1</sub> et [L <sub>1</sub> , Q <sub>1</sub> et O <sub>2</sub> ]
<b>Timbre W</b>	
Chevauchement	W <sub>1</sub>
<b>Timbre X</b>	
Non-synchronisme + A-rythmie	F <sub>2</sub> , K <sub>1</sub> , I <sub>1</sub> , P <sub>2</sub> , V <sub>1</sub> et R <sub>3</sub>
<b>Timbre Y</b>	
Synchronisme	B <sub>3</sub> , J <sub>1</sub> , O <sub>1</sub> , J <sub>2</sub>
<b>Timbre Z</b>	
Désynchronisme	A <sub>1</sub> , F <sub>2</sub> , I <sub>1</sub> , J <sub>1</sub> , J <sub>2</sub> , K <sub>1</sub> , M <sub>1</sub> , O <sub>1</sub> et R <sub>3</sub>

# Table des matières

Introduction .....	7
Chapitre 1. Le son au cinéma : des études filmiques à la narratologie audiovisuelle .....	11
Du rythme au son : Mitry et Martin .....	13
Le son chez Christian Metz .....	17
Le son dans la narratologie du cinéma .....	21
Michel Chion, ou l'art de la dénomination .....	26
Laurent Jullier et le tournant cognitif .....	30
Rick Altman et le caractère événementiel du son .....	35
Première synthèse théorique.....	39
Chapitre 2. Le rôle du son dans les études sur l'animation.....	44
Le son et la transformation de la réalité .....	47
Scott Curtis et les trois mythes de la bande sonore dans les <i>Cartoons</i> .....	52
Dessinés pour sonner (Drawn to Sound) : la « sonicité » de Rebecca Coyle .....	55
<i>Designing Sound for Animation</i> de Robin Beauchamp : vers une rhétorique du design sonore .....	61
<i>Coloriture</i> : une exploration sémiotique et narratologique du son dans l'animation.....	65
Maria Luisa Bionda : la narrativité de la musique dans le cartoon .....	70
Rita De Giuli et son compte-rendu de Jon Newsom .....	73
Giuseppe Valenzise et le pluricode de l'animation .....	74
Alberto Boschi et l'éloge du <i>Mickey-mousing</i> .....	76
Daniel Goldmark et la musique des cartoons .....	78
Deuxième synthèse théorique.....	81
Chapitre 3. L'énonciation, le corps et les dispositifs audiovisuels .....	83
Narration, monstration et énonciation .....	86
Focalisation, ocularisation, auricularisation. Le « point de vue » revisité .....	94
Auricularisation, « point d'écoute » : le corps audiovisuel.....	95
L'énonciation audiovisuelle et les modes du sensible .....	99
L'énonciation audiovisuelle comme pratique(s) sémiotique(s) .....	111
Le dispositif et l'énonciation audiovisuelle .....	117
Chapitre 4. Dispositifs audiovisuels et cinéma d'animation .....	130
Le(s) dispositif(s) animatique(s) .....	133
Le dispositif animatique du cinéma dit « expérimental » .....	139
La synthèse optique du son : un nouveau support pour de nouvelles pratiques.....	142
Norman McLaren et la pratique sémiotique du dessin sonore .....	154
La scène pratique.....	161
Les signes et les pratiques. Le clavier de papier de Norman McLaren.....	168
La recherche de la « synchronie » parfaite.....	175
Une lecture structurelle de <i>Synchronie</i> .....	179
Une structure « didactique » ? .....	179
Chapitre 5. Les sons comme événements.....	189
La cognition d'événements et l'aspectualisation des événements audiovisuels .....	193
Objets, quasi objets et événements.....	199
Les sons comme événements .....	201
Retour à la sémiotique : l'aspectualité des événements sonores .....	203
Mirai Mizue, animation abstraite et narrativité.....	203

Les timbres de Mizue : « créatures » et événements.....	211
La bande-image .....	215
La bande-son .....	224
Rythme, arythmie, a-musicalité .....	226
La combinaison audiovisuelle .....	230
Rythme, tempo, intensité : la structure tensive de la série <i>Timbre</i> .....	233
Pour finir : la série dans le cadre des pratiques sémiotiques .....	236
Conclusion.....	238
Le dispositif audiovisuel et l'actant de l'audiovision.....	246
Audiovision, visu-audition et écoute profonde .....	250
Le flux de la vie et des logiques énergétiques.....	254
Bibliographie .....	256
Filmographie .....	264
Annexes .....	272
Annexe 1 .....	272
Annexe 2 .....	280



## Table des illustrations

Figure 1 : Photogrammes du film <i>Silly Symphony: The Skeleton Dance</i> .....	8
Figure 2 : Typologie des sources sonores .....	24
Figure 3 : Des films comme textes aux films comme événements .....	36
Figure 4 : Relations du son avec les dimensions de l'expérience audiovisuelle.....	41
Figure 5 : Relations du son <i>et de l'image</i> avec les dimensions de l'expérience audiovisuelle	41
Figure 6 : Photogrammes du film <i>The Herring Murder Mystery</i> (1943) .....	80
Figure 7 : L'horizon d'invisibilité.....	106
Figure 8 : L'horizon du silence .....	106
Figure 9 : Superposition de l'audible et du visible.....	107
Figure 10 : Diagramme du temps chez Husserl .....	109
Figure 11 : Diagramme du foyer temporel.....	110
Figure 12 : Diagramme de la période temporelle.....	110
Figure 13 : Feuille à barres du film <i>Santa's Workshop</i> .....	138
Figure 14 : Le support : pellicule. <i>À la pointe de la plume</i> .....	162
Figure 15 : L'opération : dessin sur pellicule claire. Photogramme du film <i>À la pointe de la plume</i> .....	162
Figure 16 : L'opération : rayure sur pellicule noire. Photogramme du film <i>L'animateur musicien</i> .....	163
Figure 17 : Les objets-adjutants : banc et moviola. Photogramme du film <i>Experimental Animation</i> .....	163
Figure 18 : Les objets-adjutants : la moviola (image-gauche, son-droite). Photogrammes du film <i>l'animateur musicien</i> .....	164
Figure 19 : L'opérateur. Photogramme du film <i>À la pointe de la Plume</i> .....	164
Figure 20 : Les adjutants : Grant Munro (Gauche) et Evelyn Lambart (droite). Photogramme du film <i>McLaren on McLaren</i> .....	165
Figure 21 : Les objets d'inscription : les cartes. Photogramme du film <i>McLaren on McLaren</i> .....	166
Figure 22 : Les cartes. Photogramme du film <i>l'animateur musicien</i> .....	166
Figure 23 : Le banc d'animation (Mainettes et rainure). Photogramme du film <i>l'animateur musicien</i> .....	167
Figure 24 : Photogrammes du film <i>Sons sur papier</i> .....	175
Figure 25 : Photogrammes du film : <i>Tönende Handschrift II</i> .....	176
Figure 26 : Photogrammes du film : <i>Test A pour Synchronie</i> (fin des années soixante) .....	177
Figure 27 : Photogrammes du film : <i>Test B pour Synchronie</i> (fin des années soixante) .....	177
Figure 28 : La Structure générale du <i>Synchronie</i> et des photogrammes correspondants à chaque partie du film.....	180
Figure 29 : Expansion des barres (première partie) .....	181
Figure 30 : Expansion des barres (deuxième partie).....	181
Figure 31 .....	182
Figure 32 .....	185
Figure 33 : Photogrammes du film <i>Fantastic Cell</i> .....	204
Figures 34 et 35 : Phases dynamiques des sons maintenus et Phases dynamiques de la percussion-résonance .....	214
Figure 36 : Entité simple (du Timbre A).....	220
Figure 37 : Entité composée (du Timbre O) .....	221

Figures 38 et 39 : Première apparition de l'entité, Rotation90°, Rotation (45°), entité répétée et superposée .....	221
Figures 40 et 41 : Timbre K et Timbre Z .....	222
Figure 42 : <i>A long day of timbre</i> .....	223
Figure 43 : Timbre Z .....	223
Figure 44 : Photogrammes de Timbre P.....	225
Figure 45 : Photogrammes de Timbre V .....	226
Figure 46 : Photogramme de <i>Tatamp</i> .....	229
Figure 47 : Photogrammes de <i>Tatamp</i> .....	229

## Table des tableaux

Tableau 1: Les dimensions du dispositif du cinéma traditionnel .....	131
Tableau 2: Répertoire d'équivalences pour la production des sons .....	174
Tableau 3 .....	187
Tableau 4 : A et B chez Barrès (BARRES, 2006, pp. 118-119) .....	210
Tableau 5: Les entités du <i>Timbre A to Z</i> .....	216
Tableau 6: Correspondances audiovisuelles avec des timbres maintenus et avec timbres percussifs .....	230

## Résumé :

Le sujet de ce travail est le rôle du son dans le cinéma d'animation, à partir du concept d'audio-vision. Ce concept, provenant du théoricien Michel Chion, exprime une activité ou une modalité de la perception qui s'active –est qui est la face subjective– des produits audiovisuels qui dans le cas de l'animation donnent lieu à des expériences singulières. Le lieu de rencontre entre ces instances est la scène sémiotique d'une expérience de sens où une forme de textualité (résultat d'une écriture sonore et visuelle) inscrite dans un support matériel s'active et rencontre un corps vivant, qui la transforme en un récit audiovisuel, ou bien aussi en une pure expérience corporelle, rythmique ou vibratoire, qui est aussi une forme de sens. Cette scène pratique est comprise d'après le concept de dispositif, provenant de la théorie filmique, mais ici liée à la théorie de l'énonciation. Pour aborder cette scène si complexe, j'ai choisi de mettre l'accent sur un aspect particulier : le son, l'écoute, et les approches sémiotiques que se sont occupées de cette dimension de l'expérience perceptive, notamment la sémiotique du son. Pour tester cette approche, dans cette thèse on fait des analyses pratiques d'un type de productions animées que, selon le critère, on peut appeler « indépendantes », « expérimentales », ou même « abstraites » : les courts métrages animés de Norman McLaren et de Mirai Mizue. Néanmoins, pour éloignées que semblent ces productions de l'animation traditionnelle on va proposer qu'à la base de toutes ces productions il y a un même dispositif et un même type de pratique (dans le sens sémiotique du terme), un « faire technique » sur lequel on projettera des contenus divers (narratifs, dynamiques, musicaux, énergétiques).

Mots clés : Cinéma d'animation – Dispositif – Énonciation – Audio-vision – Écoute – Sémiotique – Pratique sémiotique – dispositif animatique – Son Animé – Synchronisme figural – Norman McLaren – Mirai Mizue

## **Audio-Vision in Experimental Animation. A semiotic approach**

### Abstract:

The subject of this work is the role of sound in animation, from the concept of audio-vision. This concept, from theorist Michel Chion, expresses an activity or mode of perception that is activated –and that is is the subjective side- of audiovisual productions. In the case of animation this activity gives rise to unique experiences. The meeting point between these aspects is the semiotic scene of a meaningful experience, where a form of textuality (the result of an aural and a visual writing) registered in a physical medium, is activated by a living body. This experience is interpreted as an audio-visual narrative, or as a pure physical experience of rhythm and vibration, which is also a form of sense. This practical scene is understood from the concept of 'apparatus', coined in the field of the film theory, but connected here with enunciation theory. To address this complex scene, I chose to focus on a particular aspect: sound and listening, from a semiotic perspective. To test this approach, in this work I propose practical analysis of a type of animated productions that according to the criteria, can be called "independent", "experimental", or even "abstract": the animated films by Norman McLaren and Mirai Mizue. However, even if these productions seem very distant from traditional animated cinema I propose that there is a unique type of apparatus for all animated productions, and the same type of practice (in the semiotic sense), a "technical doing" on which we project different kinds of content (narrative, dynamic musical, energy).

Keywords: Animation Film – Apparatus – Enunciation – Audio-vision – Listening – Semiotics – Semiotic practice – animatic apparatus – Animated Sound – Figural Synchronism – Norman McLaren – Mirai Mizue