

# Les sons non diégétiques dans le cinéma de fiction

Mémoire de fin d'études

**Sylvain ADAS**

Département Son, sous la direction de Jean-Pierre Laforce et Valérie De Loof  
Promotion 2020 – Abbas Kiarostami

**Juillet 2020**

# Remerciements

Je tiens en premier lieu à remercier l'ensemble des personnes dont les enseignements, l'aide et le soutien m'ont mené à l'écriture de ce mémoire.

Je remercie La Fémis, l'ensemble du personnel permanent, et particulièrement le département Son, Jean-Pierre Laforce, Valérie De Loof, Jacques Descomps, et pour cette dernière année, Tania Press, Myriam Gannagé, Barbara Turquier, ainsi que l'ensemble des intervenants dont j'ai eu la chance de profiter des enseignements au cours de mes études.

Je remercie l'ensemble des collaborateurs de mon film de fin d'études, partie pratique de ce mémoire, pour leur investissement et leur motivation sans faille, en particulier Ophélie Noury, Corentin Courage, Juliette Le Goff, Juliette Naviaux. Ce film a été une expérience si intense et si enrichissante, et il n'aurait jamais pu exister sans le concours bienveillant de toute son équipe.

Je remercie toute les personnes qui ont accepté de se prêter au jeu de l'expérience réalisée dans le cadre de la troisième partie de ce mémoire. Malgré la difficulté de l'exercice, tous l'ont mené jusqu'au bout et rien n'aurait été possible sans leur aide.

Je remercie mes parents, Nicolas et Jihane Adas, ainsi que mon frère, Jonathan, pour m'avoir toujours soutenu et pour avoir été les premiers juges de tous mes premiers balbutiements de réflexion.

Je remercie mes amis, Antoine, Nicolas, Cécile, Laura, Louise, et en cette fin d'études, mes pensées vont tout particulièrement vers mes camarades de promotion qui ont fait de ces quatre années à La Fémis une période formidable :

Claire, parce qu'il n'a jamais été aussi agréable et productif de se disputer avec quelqu'un

Victor, pour avoir été le meilleur des colocs

Sacha, pour m'avoir fait découvrir les joies et les peines, mais surtout les joies, des plans Mikoff

Noëmy, pour sa maîtrise de tous les jeux de société du monde (et pour sa relecture attentive !)

Thibaut, pour m'avoir appris la langue française

## Préambule

J'ai étudié dans ce mémoire la place qu'occupent les sons non diégétiques dans le cinéma de fiction. Cette envie est née de questionnements de spectateur avant tout. J'ai été de nombreuses fois et devant beaucoup de films très différents confronté à des façons d'utiliser des sons non diégétiques qui ont à chaque fois éveillé en moi des émotions et des sensations différentes que je ne parvenais pas à expliquer. Pourtant, le terme « diégétique » est très largement utilisé dans le langage commun de fabrication d'un film, et il m'est très souvent arrivé de parler d'une musique « qui passe de diégétique à extradiégétique » par exemple, sans réellement savoir ce qui était en jeu.

Mon approche a été de partir de la théorie la plus pure et de décortiquer avant tout le plus précisément possible de quoi on parle lorsque l'on fait référence à un « son diégétique ». Cette première étape est une tentative de définir rigoureusement l'intuition largement répandue dans le milieu du cinéma de ce qui est diégétique et de ce qui ne l'est pas. Elle était absolument nécessaire à la poursuite de la réflexion, qui ne pouvait être menée que sur des fondements solides.

Définir précisément ce qui semble évident était une tâche complexe, mais c'est à partir de cette théorie que le retour à la pratique a été possible. Une des grandes récompenses personnelles que ce mémoire m'a apportées est la possibilité de mettre en lien des façons de faire du cinéma qui n'ont *a priori* pas grand-chose à voir les unes avec les autres, mais que les outils théoriques réunissent. Les sons non diégétiques me sont alors apparus comme un élément essentiel de la mise en scène sonore au cinéma qui gagne à être étudié.

Je me suis concentré sur le « comment ? » sans traiter la question du « pourquoi ? ». Ce mémoire n'a pas pour vocation à expliquer quand utiliser un son non diégétique. J'en ai étudié la nature et les fonctions, mais la question du choix de leur utilisation mériterait une nouvelle étude.

Tout au long de ce mémoire, j'ai préféré l'expression « son non diégétique » à celle de « son extradiégétique », plus répandue. Leur sens est similaire. Dans une logique de définition, j'ai préféré laisser de côté le préfixe « extra » qui me semble orienter de lui-même la réflexion en faisant référence à quelque chose qui serait déjà connu avant d'être étudié.

# Table des matières

<b>Introduction .....</b>	<b>5</b>
<b>I. Qu'est-ce qu'un son non diégétique ? .....</b>	<b>9</b>
1. Définitions .....	9
2. Le son est invisible .....	10
3. Certaines caractéristiques des sons .....	11
a. <i>La source</i> .....	11
b. <i>Le point d'écoute</i> .....	12
c. <i>La cohérence acoustique</i> .....	13
4. Le synchronisme .....	13
5. Classification .....	14
6. Typologie des sons non diégétiques .....	15
a. <i>La source est connue mais non pertinente (la musique)</i> .....	16
b. <i>La source est connue mais c'est le langage qui apporte le sens principal (la voix parlée)</i> .....	17
c. <i>La source est connue et c'est précisément cette source qui est le sens principal (les sons réalistes)</i> .....	17
d. <i>La source est a priori inconnue mais c'est le son en tant qu'objet sonore qui transporte le sens (les sons abstraits)</i> .....	18
e. <i>Tableau récapitulatif</i> .....	18
7. Etudes de cas.....	19
a. <i>Une musique non diégétique : Star Wars</i> .....	19
b. <i>Une musique au double statut : Les Vacances de monsieur Hulot</i> .....	20
c. <i>Une voix off non diégétique : Bande à part</i> .....	21
d. <i>Une voix off ambiguë : Millennium Mambo</i> .....	23
e. <i>Un son réaliste non diégétique : The Game</i> .....	24
f. <i>Un son abstrait non diégétique : Mulholland Drive</i> .....	25
g. <i>Une source diégétique mais un point d'écoute non diégétique : Dead Man..</i>	27
h. <i>Une source diégétique et incohérence acoustique : Mulholland Drive</i> .....	28
i. <i>Un silence non diégétique : Bande à part</i> .....	29
j. <i>La diégèse ne préexiste pas au montage : Village, silenced</i> .....	30
8. Partie pratique .....	33
9. Conclusion : artificialité et caractère diégétique.....	35
<b>II. Fonctions narratives des sons non diégétiques .....</b>	<b>37</b>
1. Le grand imagier .....	38
2. Le hors-champ diégétique.....	40
3. La focalisation chez Genette.....	41
4. Configurations discursives du point de vue .....	41
5. La double temporalité .....	43
6. La distance à l'histoire : définition .....	44
7. La distance à l'histoire : critères spécifiques .....	45
a. <i>La musique</i> .....	46
b. <i>La voix parlée</i> .....	46
c. <i>Les sons réalistes</i> .....	47
d. <i>Les sons abstraits</i> .....	47
8. La distance à l'histoire : critères généraux .....	47

a.	<i>Le mixage</i> .....	47
b.	<i>La culture</i> .....	48
c.	<i>L'expérience personnelle du spectateur</i> .....	48
9.	Fonctions narratives .....	49
10.	Quelques exemples .....	50
a.	<i>Mon Oncle</i> .....	50
b.	<i>Oncle Boonmee, celui qui se souvient de ses vies antérieures</i> .....	51
c.	<i>Annie Hall</i> .....	52
11.	Partie pratique .....	53
12.	Conclusion : distance à l'histoire.....	55
<b>III. Esthétique des sons non diégétiques .....</b>		<b>56</b>
1.	Pierre Schaeffer et les trois écoutes .....	57
2.	De l'importance de l'écoute réduite au cinéma .....	58
3.	Principe et objectif de l'expérience .....	59
4.	Difficultés et limites.....	60
a.	<i>L'écoute réduite</i> .....	60
b.	<i>Limites neuroscientifiques</i> .....	61
c.	<i>Limites anthropologiques</i> .....	61
d.	<i>Limites méthodologiques</i> .....	62
e.	<i>Quel intérêt alors ?</i> .....	62
5.	Caractéristiques des sons .....	63
a.	<i>La masse</i> .....	63
b.	<i>L'allure</i> .....	63
c.	<i>L'enveloppe temporelle</i> .....	64
d.	<i>La hauteur</i> .....	64
e.	<i>La réverbération</i> .....	65
f.	<i>Le mixage</i> .....	65
6.	Protocole expérimental .....	65
7.	Analyse de la séquence .....	66
8.	Premier montage : l'imitation .....	69
9.	Second montage : la réinvention.....	71
10.	Hypothèses.....	72
11.	Résultats.....	74
12.	Interprétations des résultats .....	75
13.	Conclusion : écoute réduite et culture.....	77
<b>Conclusion.....</b>		<b>79</b>
<b>Bibliographie.....</b>		<b>82</b>
<b>Filmographie.....</b>		<b>83</b>
<b>Annexes .....</b>		<b>84</b>
(1)	Annexe 1 : séquences de l'expérience de la partie 3 .....	84
(2)	Annexe 2 : questionnaire de l'expérience de la partie 3 .....	86
(3)	Annexe 3 : résultats de l'expérience de la partie 3 .....	89

## Introduction

Le cinéma de fiction se propose de raconter une histoire en montrant des personnages agir, se déplacer, discuter, et en articulant les événements qu'ils vivent par le montage successif de différents plans. Une des conditions premières du cinéma de fiction est l'impression de réalité qu'il transmet au spectateur ; c'est-à-dire que, malgré les artifices, ce qu'il voit à l'écran se passe réellement.

Devenu sonore depuis trop longtemps pour que cela vaille encore la peine d'être précisé, le cinéma donne, la plupart du temps, à entendre ce qu'il donne à voir. En d'autres termes, il existe une homogénéité entre les images et les sons d'un film qui traduisent dans le langage du cinéma une réalité commune. Lorsqu'un personnage parle à l'écran, il est d'usage que l'on entende en même temps ce qu'il dit. C'est même la raison principale qui nous permet d'affirmer qu'il est effectivement en train de parler, et pas simplement en train de bouger ses lèvres sans émettre de son.

Pourtant, et cela avant même qu'il ne soit sonore, le cinéma a toujours été accompagné de sons qui ne faisaient pas partie de l'histoire. La « musique de film » en est l'illustration la plus connue, mais c'est aussi le cas, par exemple, du « *sound design* » de films d'horreur ou de films de science-fiction. Du côté des images par contre, la présence d'éléments extérieurs à l'histoire semble moins évidente, et en tout cas moins systématique. On peut tout de même penser à l'introduction d'archives documentaires dans un film de fiction, ou à la superposition d'animations invisibles aux personnages sur des images en prise de vue réelle par exemple.

Dans le cadre de la partie pratique de ce mémoire, j'ai réalisé un court-métrage de fiction intitulé *Summertime*. Il raconte l'histoire de Luce, une jeune femme à qui son copain, Benjamin, annonce qu'il quitte Sèvres pour toujours et qu'il ne reviendra pas. Pauline, la meilleure amie de Luce, est déjà au courant. Pour éviter qu'elle ne le retienne, ils droguent Luce qui s'endort sous le choc. Le lendemain, Luce cherche à comprendre. Elle apprend rapidement que Benjamin est censé revenir le soir même. Mais les gens autour d'elle se comportent étrangement et la réalité se distord de plus en plus.

Lors du montage et du montage son du film, nous nous sommes retrouvés confrontés avec Ophélie Noury et Victor Fleurant, respectivement monteuse et monteur son, à des

difficultés de vocabulaire quant à la désignation de certains sons. Le film possède un caractère cauchemardesque qui s'accroît au cours de l'histoire. Le traitement de la bande sonore suit cette transformation en se déréalisant petit à petit. D'abord, tout semble normal ; ensuite, l'anormalité se déguise en réalisme, mais son masque tombe peu à peu jusqu'à la séquence finale où les notions de réalisme et de son dans l'histoire ou hors de l'histoire sont tout à fait évacuées. Les difficultés de langage que nous avons rencontrées concernaient les sons que nous montions qui n'avaient pas de rapport immédiat avec les images, ou qui n'avaient pas pour cause des actions que nous voyions. La distinction classique au montage son entre les ambiances, les effets et les musiques n'était pas suffisante pour pouvoir communiquer clairement. Par exemple, nous avons monté un son continu entièrement synthétisé, qui ne s'apparente *a priori* pas à un son réaliste, mais que nous avons pensé comme un vent qui occupe l'espace de la scène. Est-il convenable de l'appeler une ambiance, rôle qu'il joue dans le montage malgré son absence de réalisme ? Est-ce plutôt une musique, en tant que nappe, malgré son absence de caractère tonal ? N'est-ce pas plutôt un effet, en référence à ce que l'on appelle couramment le *sound design*, malgré le fait qu'il soit continu et non ponctuel ?



Figure 1 *Summertime*

Il nous est apparu qu'en faisant un film qui laisse autant de place à des sons non diégétiques qu'aux sons de l'histoire, nous n'avions pas de méthode d'analyse claire qui nous permettait de savoir *a priori* pourquoi nous voulions monter certains sons et à quoi ils devaient ressembler. Cette double question de la présence d'une part et de l'apparence d'autre part du son non diégétique semble même s'opposer à l'impression

de réalité du cinéma, son artifice premier. Il est alors aisé de penser qu'un film qui briserait ce contrat d'adéquation entre les images et les sons serait perturbé dans son essence même, qu'il deviendrait une sorte de cinéma moins pur. Le montage d'un son non diégétique nécessite l'intervention consciente du réalisateur, qui décide d'ajouter quelque chose au récit, un nouvel élément. Une fois le son identifié par le spectateur comme non diégétique, il n'est pas possible de cacher le caractère volontariste de sa présence.

Mais cette difficulté apparente n'entrave pas leur utilisation, qui est très courante au cinéma. En ne prenant par exemple que le cas de la musique, il existe à toutes les époques une grande diversité de films qui l'utilisent de manières très différentes. Et mon expérience personnelle me fait remarquer que mon rapport de spectateur à ces musiques est très variable, et que cela ne dépend pas que de mon appréciation de la musique en question. La musique n'étant qu'un exemple de son non diégétique, il s'agit d'étudier dans ce mémoire la façon dont l'utilisation dans le récit de ces éléments à l'origine étrangers, presque incongrus, modifient le rapport du spectateur au film, sa réception de l'histoire.

En tant que technicien et collaborateur sonore à la création d'un film, une question s'ajoute : comment choisir ces sons non diégétiques ? Prenons par exemple le cas d'un réalisateur qui désire monter une « nappe » sur une séquence. La seule indication *a priori* que nous avons est le caractère continu du son (ce que l'on entend habituellement par « nappe »). Doit-elle être grave, ou plutôt aigue ? Doit-elle avoir un volume constant, ou être variable au cours de la séquence ? Ou encore, doit-elle être lisse, enveloppante, ou plutôt rocailleuse, granuleuse ? La première étape consiste donc à répondre à ces questions. Mais une fois cela réalisé, il existe encore une multitude de possibilités dans la création de cette nappe. A quel point cela est-il important ? A quel point la réalisation pratique de cette nappe aura une influence sur la façon dont le film sera perçu par le spectateur ? De la même façon qu'un chef opérateur doit décider de l'intensité et de l'angulation d'un projecteur, le montage de sons non diégétiques implique une multitude de choix dont il n'est pas immédiatement aisé de mesurer les implications.

Concernant leur fabrication, les sons non diégétiques possèdent une spécificité par rapport aux autres éléments de la bande sonore par leur caractère inattendu. Les sons

diégétiques sont des conséquences d'actions diégétiques, et il convient alors d'utiliser un son cohérent. Par exemple, si dans l'histoire une porte claque, on entend une porte claquer. Pour les sons non diégétiques par contre, rien dans l'histoire ne les implique *a priori*. D'une certaine façon, étant inattendus, le champ des possibles est infini. Selon quelle logique alors monte-t-on et choisit-on un son non diégétique ? Leur utilisation généralisée suggère une appartenance au langage cinématographique, et l'objet de ce mémoire est d'en examiner les règles, la grammaire.

De quelle manière l'utilisation d'un son non diégétique contribue-t-elle au récit ?

Nous définirons dans une première partie la notion de caractère diégétique d'un son, en montrant surtout qu'elle dépasse largement une distinction binaire. Nous établirons alors une typologie des sons non diégétiques, puis, par un ensemble d'études de cas, nous mettrons en évidence l'artificialité de la diégèse.

Dans une seconde partie, nous étudierons les fonctions narratives des sons non diégétiques, en faisant d'abord appel à des notions théoriques de narratologie. Cette approche nous mènera à introduire une nouvelle notion que nous avons appelé la distance à l'histoire. Combinée au caractère diégétique des sons, elle permet d'étudier le rôle qu'occupent les sons non diégétiques dans le récit filmique.

Enfin, nous réaliserons dans une troisième partie une expérience à partir d'une séquence extraite du film *Mulholland Drive* de David Lynch. Après l'avoir justifiée par une présentation des différents types d'écoute au cinéma, elle nous permettra d'étudier l'importance de l'esthétique des sons non diégétiques, et ce indépendamment de leurs fonctions narratives, de la place qu'ils occupent dans le montage.

## I. Qu'est-ce qu'un son non diégétique ?

La question de l'utilisation des sons non diégétiques dans le cinéma de fiction requiert une définition convenable de la diégèse du film et de l'appartenance ou la non appartenance d'un son à celle-ci. Il s'agit simplement de préciser le cadre d'étude de ce mémoire, c'est-à-dire de déterminer lors de l'analyse d'une séquence si un son est ou n'est pas diégétique.

### 1. Définitions

Le terme diégèse est employé avec deux sens différents. Le premier est le fait de raconter les choses. Pour Platon, l'imitation (*mimèsis*) est opposée au récit (*diègèsis*) ; la monstration en opposition à la narration. Aristote considère quant à lui le récit (*diègèsis*) comme un mode de l'imitation poétique (*mimèsis*), là aussi en opposition à la représentation (le théâtre)<sup>1</sup>.

Le second sens du mot « diégèse » a été établi par Étienne Souriau dans la *Revue internationale de filmologie* en 1951 : « tout ce qui est censé se passer, selon la fiction que présente le film ; tout ce que cette fiction impliquerait si on la supposait vraie. » C'est avec ce dernier sens que nous utiliserons ce terme, c'est-à-dire l'univers du film.

Ainsi, un élément du film est dit « diégétique » lorsqu'il appartient au monde dans lequel les actions de l'histoire ont lieu. Par exemple, le personnage principal d'une fiction est diégétique ; par contre, l'acteur qui l'incarne ne l'est pas.

Deux espaces différents et opposés sont mis en évidence par cette définition. D'un côté, il y a le monde de l'histoire, celui dans lequel se déroulent les actions mises en scène par le film, et de l'autre, tout ce qui n'en fait pas partie. Cet autre côté, c'est d'une part ce que l'on appelle le monde réel, mais c'est également celui de toutes les autres diégèses. Ainsi, le gag de bande dessinée où des personnages d'un autre univers se trompent de case et interviennent dans une histoire à laquelle ils n'appartiennent pas est un mécanisme extradiégétique.

---

<sup>1</sup> Genette Gérard. Frontières du récit. In: Communications, 8, 1966. Recherches sémiologiques : l'analyse structurale du récit. pp. 152-163. ([https://www.persee.fr/doc/comm\\_0588-8018\\_1966\\_num\\_8\\_1\\_1121](https://www.persee.fr/doc/comm_0588-8018_1966_num_8_1_1121))

Dans la diégèse d'un film de fiction, la caméra, la plupart du temps, n'existe pas. C'est donc le même genre de procédé qui entre en jeu lorsqu'un personnage s'adresse face caméra directement au spectateur, lorsqu'il « brise le quatrième mur ». Cette action du personnage met en lien deux univers hétérogènes, et redéfinit par son existence même la diégèse du film.

Il est important de remarquer que dire d'un élément qu'il est diégétique ou non diégétique relève nécessairement de la subjectivité du spectateur. Discriminer les éléments diégétiques et non diégétiques d'un film requiert une confrontation de l'œuvre aux connaissances et à l'expérience de chacun. Cela sera particulièrement important dans le cadre des sons, mais est observable bien plus largement. Par exemple, lorsque l'on dit d'un acteur qu'il « joue pour la caméra », sous-entendu il n'est pas vraiment « dans » son personnage, on considère que certains éléments de son comportement ne sont pas diégétiques ; cela relève de l'interprétation personnelle du spectateur.

## **2. Le son est invisible**

Qu'en est-il alors des sons ? Contrairement à ces deux derniers exemples, dans la grande majorité des cas, les sons non diégétiques n'interagissent pas avec l'univers du film. Les personnages n'entendent pas la « bande originale », la musique du film. Michel Chion appelle cette musique « de fosse », au sens où elle pourrait être jouée par un orchestre dans la fosse ; elle est audible par les spectateurs du film mais pas par les personnages.

Nous avons immédiatement simplifié cette question de l'interaction entre un son non diégétique et l'univers du film par le fait qu'un personnage l'entende ou ne l'entende pas. Cela mérite d'être précisé.

Le son, par son caractère diffus et non localisé (sa source l'est, mais pas le son en lui-même) et par la faible énergie qu'il transporte, n'interagit pas mécaniquement avec ce qu'il traverse, ou en tout cas, cela n'est pas perceptible à l'œil nu. Par exemple, un son résonant dans une pièce où du linge est étendu ne va pas faire bouger ce linge. Au plus, il fera trembler la membrane d'une enceinte, mais là encore, le mouvement est presque imperceptible.

Cette caractéristique du son a une conséquence importante : on ne peut pas *voir* le son. On peut observer sa cause lorsqu'elle est visible, on peut constater la transformation de sa texture par la position relative de sa source à son auditeur et par l'acoustique d'une pièce, mais on ne peut pas le voir évoluer. Les sons diégétiques et non diégétiques ont donc en commun leur invisibilité. On ne peut pas dire d'un son qu'il est diégétique parce qu'on le voit, et qu'il n'est pas diégétique parce qu'on ne le voit pas.

Ainsi, l'interaction entre un son et la diégèse n'est pas évidente *a priori*. Il convient donc de déterminer ce qui permet de discriminer les sons diégétiques et non diégétiques.

### **3. Certaines caractéristiques des sons**

#### *a. La source*

Le premier élément qui permet de déterminer le caractère diégétique d'un son est la présence dans le film de sa source. L'attitude d'écoute causale dans laquelle se trouve le spectateur d'un film permet l'association des événements, des objets, des personnages dans l'histoire avec les sons qu'ils émettent. C'est bien sûr l'expérience de chacun qui permet d'associer un son à sa source (comme une sonnerie de téléphone au combiné), mais c'est également le synchronisme qui permet d'orienter cette association, notamment pour des sons peu courants (comme des machines complexes dans des usines par exemple).

Il n'est pas nécessaire que cette source soit visible dans le champ. Il est d'ailleurs même rare qu'elle le soit lorsqu'il s'agit de sons d'ambiance par exemple, comme des oiseaux dans des arbres. Là encore, c'est l'expérience de chacun qui permet d'associer ce qu'il entend au lieu qu'il voit, grâce au contexte de l'histoire. Rien n'assure que les oiseaux soient effectivement présents dans les arbres, mais lorsque l'expérience personnelle du spectateur rend cette association possible, il est tout à fait acceptable de dire que ce son est dans le film.

On pourra donc dire d'un son que sa source est diégétique lorsqu'elle est présente, ou *probablement* présente, dans l'histoire, et qu'elle n'est pas diégétique dans le cas contraire.

b. *Le point d'écoute*

Selon la définition de Michel Chion, on appellera point d'écoute d'un son la personne dont l'écoute pourrait correspondre à celle des sons de la bande sonore du film ; autrement dit, à la place de qui le spectateur écoute.

Chion précise cette notion en distinguant d'une part le point d'écoute spatial, qui correspond à celui de la position hypothétique d'un micro dans le lieu, c'est-à-dire ce que nous entendrions si nous étions à la place de ce micro ; la plupart du temps il correspond à celui de la caméra, mais le cas classique de personnages loin de la caméra que l'on entend parler clairement est un exemple d'un autre point d'écoute spatial (c'est ce que Laurent Jullier appelle une « scission Point de Vue – Point d'Ecoute »<sup>2</sup>). Il distingue d'autre part le point d'écoute subjectif, qui correspond au personnage par les oreilles duquel il est suggéré que nous entendions un son. Ce point d'écoute subjectif permet de prendre en charge certains sons à l'analyse complexe, comme une voix off d'un personnage-narrateur par exemple.

Dans le cadre de ce mémoire, contrairement à l'analyse de Michel Chion, on n'associera pas la notion de point d'écoute à une séquence audio-visuelle dans sa globalité mais plutôt à chaque son de cette séquence, chacun pouvant correspondre à un point d'écoute différent. Il est important de noter que le point d'écoute d'un son peut être à la fois spatial et subjectif ; c'est le cas pour un son dont la source est diégétique et qu'on entend comme un personnage qui est à l'écran.

On peut remarquer que le point d'écoute spatial n'a de sens que lorsque la source du son est diégétique.

On dira donc que le point d'écoute d'un son est diégétique lorsqu'il est :

- Spatial, c'est-à-dire que sa source est diégétique et qu'on peut déterminer depuis où dans l'espace du film on l'entend ;
- Subjectif, c'est-à-dire qu'il est suggéré qu'un personnage du film l'entend, même si sa source n'est pas diégétique.

Cela permet donc de mettre en évidence des cas de points d'écoute non diégétiques, comme la musique de fosse orchestrale par exemple.

---

<sup>2</sup> *Le son au cinéma*, Laurent Jullier, p. 56

### c. *La cohérence acoustique*

On dira d'un son qu'il est cohérent acoustiquement lorsque sa source est diégétique et que son apparence est suffisamment crédible. Cela concerne trois aspects :

- La concordance à la source : est-ce que le son ressemble suffisamment à ce à quoi on s'attend ?
- La réverbération : est-ce que la façon dont ce son résonne correspond à l'espace dans lequel il évolue ?
- La cohérence spatiale : est-ce que le point d'écoute spatial est possible ?

Il s'agit ici de déterminer ce qui rend un son *vraisemblable*. Cela relève de la perception personnelle du spectateur, et particulièrement de la conformité du son qu'il reçoit avec l'image qu'il se fait du monde réel<sup>3</sup>. Ainsi, il est nécessaire de savoir qu'un téléphone sonne pour remarquer qu'il ne sonne pas normalement ; de la même façon, il faut être entré dans une église pour savoir comment les sons résonnent à l'intérieur. La cohérence spatiale relève quant à elle moins des connaissances du spectateur que de son expérience sensible du monde ; par exemple, il n'est pas vraisemblable que deux personnes discutent en chuchotant tout en étant dos à dos, ou loin l'une de l'autre.

## 4. **Le synchronisme**

La question du synchronisme est intimement liée à celle de la diégèse. Par synchronisme, on n'entend pas ici simplement le rapport entre une action dans l'image et le montage au même moment d'un son qui s'y réfère. La question est plus globale, et relève surtout du lien sémantique qui peut exister, ou non, entre la bande image et la bande sonore. *Est-ce que l'histoire racontée par les images et l'histoire racontée par les sons est la même ?*

C'est pourquoi ce sujet se limite au cinéma de fiction, c'est-à-dire les films qui tentent par le montage simultané d'images et de sons de raconter une histoire. Si l'on prend un film dont on retire le son, et qu'on y ajoute celui d'un autre film, on aura toujours un objet audio-visuel difforme, dans lequel sans aucun doute apparaîtront au hasard des

---

<sup>3</sup> 200 mots-clés de la théorie du cinéma, André Gardies, Jean Bessalet, p. 210

points de synchronisation (au sens de Michel Chion), mais duquel aucun synchronisme fort ne peut se dégager.

Cela est peut-être arbitraire. Il est tout à fait possible d'imaginer un film où les images et les sons présentent un asynchronisme apparent, mais qui aurait tout de même l'ambition de raconter une histoire commune. Acceptons cette limite, et ne considérons que les films où l'image et le son nous semblent, à titre personnel, évidemment synchrones. Laurent Jullier évoque ce cas de mariage parfait :

*« Il arrive que nous ayons l'impression que bande-son et bande-image collent merveilleusement bien l'une avec l'autre, et qu'il ne saurait être question d'imaginer ces deux images-là avec d'autres sons ou l'inverse – cette « correspondance » parfaite, chère aux Romantiques, s'appelle un effet synesthétique ; c'est un moment où le terme même d'audiovisuel ne peut s'écrire qu'en un seul mot, sans trait d'union aucun. David Lynch y est très sensible : « Tout travaille dans le même sens, et ça vous transporte... Quand ça marche, c'est quelque chose de magique, qui peut vous envoyer, vous savez, plus haut » (Soundscape, 2003).<sup>4</sup> »*

## 5. Classification

Dans *Le son au cinéma*<sup>5</sup>, Laurent Jullier propose une classification des « mondes » auxquels appartiennent les sons d'un film :

- le **monde diégétique** : celui de l'histoire, le monde dans lequel vivent les personnages.
- les **mondes intérieurs** : ce à quoi les personnages sont susceptibles de penser ; ce sont des sons dont le point d'écoute est subjectif, et dont la source n'est pas forcément diégétique. Laurent Jullier insiste sur le fait que parfois, il ne relève que du spectateur de décider qu'un son appartient au monde diégétique ou à un monde intérieur.
- les **fosses imaginaires** : en analogie avec le théâtre, l'endroit qui accueille l'orchestre, la voix off d'un narrateur, etc.

---

<sup>4</sup> *Le son au cinéma*, Laurent Jullier, p. 59-60

<sup>5</sup> *ibid.*, p. 56-58

- le monde de la **production matérielle** : le plus distant avec le monde diégétique, celui qui n'apparaît dans le film qu'au générique par exemple, ou par des erreurs techniques comme un micro dans le champ ou un rail de travelling dans un coin de l'image.

Cette classification permet de s'éloigner de la distinction binaire entre le diégétique et le non diégétique. Mais ce qui complexifie le sujet, c'est que ces *Combinaisons Audio-Visuelles* (CAV), c'est-à-dire les rapports entre la bande image et la bande son, ne sont pas stables. C'est ce que Laurent Jullier appelle les « *métamorphoses des CAV* ». Ces métamorphoses suivent en général deux tendances, soit à s'éloigner du monde diégétique, soit à s'en rapprocher. Les exemples sont très nombreux, et nous aurons l'occasion plus tard lors d'études de cas de les décrire. En bref, le caractère diégétique d'un son peut évoluer dans le temps, il n'est pas stable.

Cette classification possède cependant une limite. Comme nous l'avons évoqué plus tôt, certains sons n'appartiennent pas tout entier à un monde ou à un autre, mais ils possèdent certaines caractéristiques qui peuvent être classées dans ces catégories. Ainsi, non seulement un son peut passer d'un monde à un autre au cours de son existence dans le film, mais il peut également appartenir à plusieurs de ces mondes à la fois.

Nous utiliserons donc cette classification bien pratique pour parler précisément du niveau de diégèse d'un son, tout en gardant en tête qu'elle ne peut être universelle. Trop d'exemples la contrediront, mais c'est précisément par ces contradictions que ces exemples pourront être analysés (un son de la production matérielle tout à coup diégétique par exemple).

## 6. Typologie des sons non diégétiques

Maintenant que nous avons mis en évidence une façon de déterminer le caractère diégétique d'un son, nous proposons de les distinguer en quatre catégories différentes.

Il est traditionnel de classer les éléments d'un montage son en trois grandes familles :

- Les ambiances : sons continus, caractérisent l'espace ; nous pouvons préciser dans le cadre de ce mémoire qu'ils sont toujours diégétiques.

- Les effets : sons surtout ponctuels, ce qui n'est pas présent dans le son direct ; diégétiques ou non diégétiques.
- Les musiques : originales ou préexistantes ; diégétiques ou non.

Nous le voyons, cette classification n'est pas adaptée à une analyse claire de leur caractère diégétique. Afin de dépasser cette limite, remarquons qu'un son au sein d'un montage possède deux caractéristiques fondamentales :

- Le sens (c'est-à-dire le fond) ;
- L'aspect (c'est-à-dire la forme).

Les trois catégories précédentes, ambiances, effets et musiques, discriminent les sons surtout selon leur aspect (le son est-il continu, ponctuel, musical). Pour traiter le cas des sons non diégétiques, nous proposons de les catégoriser non pas selon leur sens immédiatement, mais plutôt selon ce qui transporte le sens, selon l'endroit où il est contenu. Cela a à voir avec la source du son.

*a. La source est connue mais non pertinente (la musique)*

Cette première catégorie regroupe les sons dont la source est connue mais la connaître ne nous renseigne pas sur le sens qu'elle véhicule. Elle s'apparente la plupart du temps à la musique.

La musique est le type de sons non diégétiques le plus courant dans le cinéma de fiction. Il peut s'agir d'une musique originale, composée pour le film, ou d'une musique préexistante. Le choix de la musique est déterminé par des questions culturelles et esthétiques, et son utilisation est très variable.

Dans certains cas, la distinction entre une musique et d'autres types de sons non diégétiques est délicate. C'est pourquoi les regrouper selon ce rapport à leurs sources permet de les distinguer précisément.

Ainsi, lorsque l'orchestre ouvre *Star Wars* (George Lucas, 1977), ce n'est pas le fait que ce soit un orchestre qui est pertinent, c'est la musique en elle-même. De la même façon, lorsque les gouttes d'eau dans *Bambi* (David D. Hand, 1942) deviennent musicales, ce

n'est plus le fait qu'elles soient des gouttes d'eau qui est pertinent, c'est la musique qu'elles créent.

*b. La source est connue mais c'est le langage qui apporte le sens principal (la voix parlée)*

La voix parlée non diégétique, dont l'exemple le plus immédiat est la voix off, est un type très courant de son non diégétique dans le cinéma de fiction. Elle peut appartenir soit à un personnage du film soit à quelqu'un qui y est extérieur, comme un narrateur dont on ne sait rien ou bien le réalisateur même du film.

Son traitement est souvent de bonne qualité technique, enregistrée dans un espace peu réverbérant et mixée suffisamment fort pour qu'elle se détache bien du reste de la bande sonore et que sa compréhension soit simple. Cependant, elle peut également être tout à fait différente, comme dans le cas d'utilisation d'archives sonores documentaires par exemple.

Dans ce cas, la source est connue (des cordes vocales), mais le sens est avant tout contenu dans les mots prononcés.

*c. La source est connue et c'est précisément cette source qui est le sens principal (les sons réalistes)*

Nous proposons de qualifier de « réalistes » les sons non diégétiques autres que la musique et la voix parlée dont la source est identifiable par le spectateur sans qu'il la voie. En d'autres termes, ce sont des sons qui pourraient être diégétiques si leur source appartenait au monde de l'histoire.

Parfois, la distinction avec la catégorie suivante, celle des sons « abstraits », est délicate. Ce que l'on peut appeler « *l'identification acousmatique* », c'est-à-dire le processus actif du spectateur d'identification de la source invisible d'un son, est fondamentalement subjective. De ce fait, ce qu'un spectateur pourrait associer à une certaine source ne serait pas nécessairement perçu de la même façon par quelqu'un

d'autre, relativement à l'expérience personnelle de chacun et à la façon dont le film est reçu.

*d. La source est a priori inconnue mais c'est le son en tant qu'objet sonore qui transporte le sens (les sons abstraits)*

La dernière catégorie de sons non diégétiques correspond à tout le reste, que nous appellerons « abstraits ». Ce sont des sons dont l'écoute causale ne mène à rien, et qui gardent pour eux le mystère de leur source. La situation d'écoute du spectateur est une écoute réduite, au sens de Pierre Schaeffer.

C'est par exemple le cas de ce à quoi on se réfère lorsque l'on parle de « *sound design* », comme dans un film de science-fiction ou un film d'horreur par exemple. La distinction avec la musique et les sons réalistes peut être ardue, et c'est pourquoi elle dépend de la capacité du spectateur à identifier ou non la source du son.

*e. Tableau récapitulatif*

Le tableau suivant récapitule ces quatre catégories de sons non diégétiques :

<b>Catégorie</b>	<b>La source est connue</b>	<b>La source est pertinente</b>	<b>Sens</b>
<b>Musique</b>	Oui	Non	Musical
<b>Voix parlée</b>	Oui	Non	Verbal
<b>Sons réalistes</b>	Oui	Oui	Causal
<b>Sons abstraits</b>	Non	Non	Ecoute réduite

Nous proposons de se référer dans la suite de ce mémoire à ces quatre catégories par les termes « musique », « voix parlée », « sons réalistes » et « sons abstraits », afin de simplifier le discours. Il est toutefois important de ne pas oublier que la voix parlée peut être musicale, que la musique peut être montée en tant que référence culturelle (son sens principal est alors causal), etc. Nous mettrons en évidence de nombreux cas particuliers et certains sons non diégétiques appartiendront à plusieurs catégories à la fois, mais ces

limites ne justifient pas selon nous de ne pas simplifier la dénomination de ces catégories.

## 7. Etudes de cas

Afin de mettre en application les définitions développées dans cette partie, nous traitons une série de séquences contenant des sons non diégétiques ou à caractère non diégétique, d'abord simples et quelques cas moins évidents ensuite.

Plutôt qu'une forme de catalogue, il s'agit de choisir des exemples qui permettent de mettre en évidence la diversité des sons non diégétiques utilisés dans le cinéma de fiction. A chaque fois, on tentera de répondre à la question « est-ce que ce son est diégétique ? ».

Au travers de ces exemples, nous souhaitons montrer que la logique qui régit le montage de sons non diégétiques est purement cinématographique et n'est pas subordonnée à la logique du monde réel.

### a. Une musique non diégétique : *Star Wars*

Dans *Star Wars* (George Lucas, 1977), comme dans tous les autres films qui l'ont suivi dans le même univers, la musique composée par John Williams occupe une place prédominante. Suivant le concept du *leitmotiv*, elle accompagne le film du début à la fin, et en est un élément mémorable. Est-elle diégétique ?



Figure 2 *Star Wars*

Evidemment, la source de ces sons, c'est-à-dire l'orchestre, n'est pas diégétique : il n'appartient pas à l'univers du film. La question du point d'écoute spatial ne se pose donc pas, de la même façon que celle de la cohérence acoustique.

Cette musique est-elle le fruit d'un point d'écoute subjectif ? Autrement dit, est-ce qu'un personnage l'entend dans sa tête ? Cela ne semble pas du tout être le cas.

Ainsi, cette musique est tout à fait non diégétique. Elle appartient au monde des *fosses imaginaires* (selon la classification de Laurent Jullier), et au type *musique* de sons non diégétiques que nous avons définis plus tôt, c'est-à-dire que sa source (l'orchestre) est connue mais non pertinente.

*b. Une musique au double statut : Les Vacances de monsieur Hulot*

*Les Vacances de monsieur Hulot* (Jacques Tati, 1953) présente un cas intéressant de trouble diégétique lors d'une séquence où une vacancière place un disque sur un tourne-disque et lance la musique. C'est encore le début du film, mais on a déjà entendu plusieurs fois le thème composé par Alain Romans. La musique de la vacancière démarre alors, mais c'est la musique du film qui joue.

Comment ce personnage peut-il connaître la musique du film auquel il appartient ? D'un point de vue analytique, la source du son est *a priori* diégétique. Cependant, la cohérence spatiale est ici perturbée. Que l'on soit à l'intérieur de l'hôtel ou à l'extérieur lorsqu'elle va à la fenêtre, on entend la musique de la même façon. Puis le film la quitte et rejoint d'autres vacanciers, et la musique reste la même : elle joue cette fois le rôle de musique de fosse, aussi peu diégétique que la musique de *Star Wars* que nous avons traitée ci-dessus.

Ainsi, la caractéristique principale de cette musique qui trouble son rapport à la diégèse est la concordance à la source. On s'attend à ce qu'une musique sorte du tourne-disque, et c'est bien une musique qui démarre au même moment. Cependant, celle-ci possède toutes les caractéristiques d'une musique non diégétique, et il est déjà connu en l'entendant que c'est le thème du film. En fait, il est probable qu'on n'entende pas du tout la même musique que la vacancière. Ce retour à la musique du film crée une distance avec l'action, une sorte de retour à la narration globale qui se détache de ce

personnage en particulier. Nous analyserons ces mécanismes plus en détail dans la seconde partie de ce mémoire.



Figure 3 *Les Vacances de monsieur Hulot*

c. *Une voix off non diégétique : Bande à part*

La scène où les trois personnages dansent dans *Bande à part* (Jean-Luc Godard, 1964) présente un cas de voix off clairement non diégétique.

Au milieu du café bruyant, perturbés par les serveurs qui continuent leur travail, les trois personnages suivent la même chorégraphie au rythme de la musique. Tout à coup, la musique (et les bruits du restaurant ?) se coupent, et seuls restent le son des pas des personnages et de leurs claquements de doigts. La voix de Jean-Luc Godard annonce alors que « il est temps d'ouvrir une deuxième parenthèse et de décrire les sentiments des personnages », ceux-ci continuant à danser, indifférents au brutal changement de ton de la scène. Puis la musique reprend, comme si rien ne s'était passé, avant de se recouper une nouvelle fois pour laisser place à la voix off, toujours accompagnée des pas et des claquements de doigts des personnages ; ainsi de suite sur le reste de la séquence.



Figure 4 *Bande à part*

Que l'on reconnaisse ou pas la voix de Jean-Luc Godard, donc du réalisateur, cela n'a pas vraiment d'importance ; il est de toute façon clair que ce n'est pas la voix de quelqu'un dans l'histoire. Ainsi, la source du son n'appartient pas à l'histoire. Elle s'adresse au spectateur et n'est entendue par aucun personnage. Cette voix parlée est donc non diégétique, son sens principal est contenu dans le langage, et elle appartient au monde des *fosses imaginaires*. Cette voix off est celle d'un narrateur omniscient ; il semble même que c'est cette même personne qui fait cesser par sa simple volonté (celle de faire une parenthèse) le reste des sons de la séquence.

Mais alors, pourquoi persistent les sons de pas et de claquements de doigts ? Alors qu'ils appartiennent au même niveau diégétique que la musique, eux ne cessent pas. Le spectateur initié comprend alors que ces sons proviennent du bruitage, et qu'ils sont donc fabriqués. Tout à coup, c'est le dispositif même du cinéma qui est révélé par cette parenthèse. Cela met même en péril le caractère diégétique de ces sons qui semblaient produits par les personnages : non, ils n'entendent pas le bruit de leurs pas en dansant puisqu'ils proviennent du bruitage.

Pire encore, la séquence se termine sur le personnage interprété par Anna Karina resté seul sur la piste de danse et qui demande « où est Arthur ? », et la musique s'interrompt au moment même où elle commence à parler. Une fois de plus, le dispositif

cinématographique est révélé : la musique est montée, les personnages ne l'entendent pas.

Pourtant, nous préférons toujours considérer que ces sons de pas, de claquements de doigts et la musique sont diégétiques. Que ces effets de montage révèlent le dispositif cinématographique est une chose. Cependant, ce que tout cela permet est de raconter l'histoire que ces personnes dansent sur de la musique dans un café en se posant des questions, plus ou moins pragmatiques et plus ou moins existentielles. Ainsi, nous sommes d'avis de considérer que malgré ces difficultés, il s'agit quoi qu'il en soit de raconter une histoire grâce au langage cinématographique. La question de la diégèse se distingue donc d'une analyse filmique cartésienne ; elle n'est pas nécessairement mise en péril par la dissociation entre l'image et le son, ou plus généralement par la révélation que le film est une fabrication.

*d. Une voix off ambiguë : Millennium Mambo*

La séquence d'ouverture de *Millennium Mambo* (Hou Hsiao-hsien, 2001) suit une femme, le personnage principal du film interprété par Shu Qi, marchant dans un long tunnel éclairé par des néons. La bande sonore est constituée de deux éléments seulement : une voix off et une musique. Pour les mêmes raisons que celles exposées plus tôt, la musique n'est pas diégétique. Concentrons-nous sur la voix off.



**Figure 5** *Millennium Mambo*

Cette voix raconte l'histoire d'une rupture amoureuse impossible et les projets de la femme de s'en aller pour de bon. Qu'on le comprenne sur le coup ou *a posteriori* en regardant le film, cette voix est celle de la femme que l'on suit, et celle-là même de l'histoire. Cependant, elle parle d'elle-même à la troisième personne, et selon ses dires, dix ans après les événements qu'elle décrit. Même si la source de ce son n'est pas immédiatement dans l'image (elle ne dit pas ces paroles tout en marchant), elle est tout de même diégétique car la personne qui la dit appartient à l'histoire.

C'est une voix qui possède toutes les caractéristiques d'une voix off non diégétique, si ce n'est le fait que la personne qui la dit est diégétique. Peut-être provient-elle de son point d'écoute subjectif ? Probablement pas : elle est adressée au spectateur, ou en tout cas à quelqu'un qui n'est pas dans l'histoire.

Le doute sur son caractère diégétique provient de cette adresse au spectateur, à cet élément étranger à l'histoire. Elle ne semble exister que parce que le film existe, comme si elle prenait conscience de son caractère filmique.

Selon les types de sons non diégétiques que nous avons décrits, elle appartient à la voix parlée : le sens est contenu dans le langage. Cependant, elle entre dans le monde diégétique parce que sa source est diégétique. Cette source est donc pertinente dans son appréciation (ce n'est pas un narrateur neutre) : elle est tout autant porteuse de sens que le langage.

*e. Un son réaliste non diégétique : The Game*

Dans *The Game* (David Fincher, 1997), le personnage principal incarné par Michael Douglas voit sa vie bouleversée par un jeu de rôle organisé par une mystérieuse entreprise. Un soir, il rentre chez lui et trouve une marionnette de clown à taille humaine devant sa porte. Il la traîne alors à l'intérieur et l'ouverture de la porte est marquée par un son de coup de feu.



Figure 6 *The Game*

La source de ce coup de feu n'est pas diégétique. Le point d'écoute ne semble pas non plus subjectif. Bref, ce son n'est pas diégétique.

Il semble être monté à la place du son d'ouverture de porte, enfoncée violemment par le personnage angoissé. En tout cas, cette marionnette traînée au sol ressemble tout à fait à un cadavre. Ce son de coup de feu appartient lui aussi à ce thème du meurtre.

Il appartient donc à la catégorie des sons réalistes : la source est connue (le coup de feu) et c'est elle qui est porteuse du sens (elle évoque le meurtre). Il appartient au monde des *fosses imaginaires* au même titre que la musique de film, mais il s'en distingue par le rapport qu'a le spectateur avec sa source.

f. *Un son abstrait non diégétique : Mulholland Drive*

Dans *Mulholland Drive* (David Lynch, 2001), la bande sonore contient un enjeu de narration majeur. La première des deux parties du film est un rêve du personnage principal, interprété par Naomi Watts, et le dispositif cinématographique même adopte ce point de vue subjectif. C'est-à-dire que chaque mouvement de caméra, chaque son entendu, et même chaque scène vue, est le résultat de la pensée du personnage endormi.

Ainsi, la nature onirique du film est propice à un éloignement d'un certain réalisme cinématographique, et donc à l'utilisation de sons non diégétiques. Ceux-ci sont de diverses natures, et nous y reviendrons plusieurs fois ; pour l'instant, concentrons-nous sur une scène particulière qui propose une autre utilisation de sons abstraits.

Vers la fin de la première partie du film, les deux personnages principaux se rendent au milieu de la nuit au club *Silencio*. Elles marchent d'abord dans la rue et prennent un taxi, et l'image devient floue et la caméra vibre brusquement jusqu'à la coupe. Puis le taxi les conduit à travers Los Angeles dans la nuit, et se succèdent les images défilantes de la ville à travers les vitres, jusqu'à ce qu'il les dépose à leur destination, sur une sorte de parking où des déchets remuent au vent. La caméra plonge alors dans un brutal travelling avant vers l'entrée du *Silencio*.

Les mouvements de caméra du début et de la fin de la séquence sont à chaque fois accompagnés de sons abstraits synchrones. Mais ceux qui nous intéressent sont ceux accompagnant la scène dans la voiture.

D'abord, on entend un moteur plutôt réaliste. Mais rapidement, celui-ci se transforme en un son de même nature, grave, continu, mais n'ayant plus grand-chose à voir avec ce moteur. Ensuite, les coupes sur de nouveaux plans à travers les vitres sont accompagnées de sons plus aigus, presque venteux, mais n'ayant eux non plus grand-chose à voir avec du vent. Enfin, lorsque la voiture arrive devant le *Silencio*, un autre son venteux, devenant presque musical, comme des cuivres peut-être, prend le relai.

En passant de nouveau la scène mais les yeux fermés cette fois, le ressenti est totalement différent. On ne sent plus du tout les coupes, mais une sorte de son à la forme complexe évoluant tout au long de la séquence. On ne voit plus de voiture, on ne voit plus de vent, et on ne sent plus le passage à l'extérieur de la fin de la séquence.

Ainsi, un phénomène de tentative d'identification de la source de ces sons abstraits se produit en regardant la scène, qu'on pourrait qualifier d'identification acousmatique. Etant incapables de rattacher ces sons à quelque chose de concret, mais ceux-ci étant apparemment liés à des événements visuels, nous les associons à des causes auxquelles ils sont en fait totalement étrangers.

Il s'agit ici d'un cas où la source du son n'est pas diégétique, mais où l'association synchrone entre ce son abstrait et un événement diégétique crée un rapport de causalité entre eux. Ce n'est finalement pas si différent de l'exemple des *Vacances de monsieur Hulot*, où la musique non diégétique était déclenchée par la vacancière. Pourtant, il ne

s'agit pas ici de musique, parce qu'à la différence du film de Tati, les sons utilisés ici gardent pour eux le mystère de leur source.

C'est une conséquence du phénomène que Michel Chion appelle la synchrèse : « soudure irrésistible et spontanée qui se produit entre un phénomène sonore et un phénomène visuel ponctuel lorsque ceux-ci tombent en même temps, cela indépendamment de toute logique rationnelle. »

*g. Une source diégétique mais un point d'écoute non diégétique :  
Dead Man*

La scène d'introduction de *Dead Man* (Jim Jarmusch, 1995) présente le personnage principal incarné par Johnny Depp se rendant dans la ville de Machine, le terminus du train qu'il emprunte. La séquence est divisée en plusieurs segments, séparés par des fondus au noir. A chaque fois, les voyageurs du train changent, partant d'une population familiale pour arriver aux rebuts de la société.



**Figure 7** *Dead Man*

De nombreux types de plans sont utilisés à chaque itération du train, créant une logique de répétition, démarrant par les roues accompagnées de la musique de Neil Young, puis entrant dans le train en suivant le personnage observant ses compagnons de route, lisant, tentant de dormir, etc. Un de ces plans est celui du regard du personnage vers l'extérieur

du train, où il voit le paysage défiler et évoluer au cours du trajet, la fenêtre en amorce. A chaque fois, sur ce plan, on peut entendre le vent siffler.

La source de ce son de vent est dans l'image, mais ce qui pose problème au niveau diégétique est le point d'écoute. La fenêtre étant fermée, le personnage à l'intérieur, le fait que l'on entende l'extérieur uniquement au moment où on le voit est une fabrication cinématographique qui n'a pas de logique *a priori*. C'est un exemple de scission point de vue – point d'écoute.

En découvrant la séquence, ce son ne pose pas question. Il semble logique de l'entendre. De la même façon que la musique non diégétique est associée aux roues du train, ce son à caractère non diégétique est associé au paysage qui défile. Ainsi, malgré l'absence de logique au sein de la diégèse, le montage de ce son répond à une logique au sein du dispositif cinématographique.

*h. Une source diégétique et incohérence acoustique : Mulholland Drive*

Retour à *Mulholland Drive*. Parmi les premières scènes du film, une d'entre elles montre des personnages mystérieux dont on ne voit pas grand-chose se passer des coups de fil. Le premier indique au second que « La fille est toujours portée disparue. » Le second raccroche, compose un numéro ; un téléphone sonne et un troisième homme, dont on ne voit que le bras, répond. Le second lui indique : « La même chose. » Ce troisième homme raccroche également et compose lui aussi un numéro. Un téléphone noir sonne, une fois, deux fois, puis la troisième fois résonne magiquement jusqu'à la séquence suivante et introduit la musique accompagnant l'arrivée du personnage principal à l'aéroport de Los Angeles.



Figure 8 *Mulholland Drive*

Cette dernière sonnerie de téléphone, alors qu'elle était jusqu'ici diégétique, présente une incohérence acoustique. Elle n'est pas vraisemblable.

Il est intéressant de remarquer que les dialogues mêmes de cette séquence présentent un problème diégétique. Comment le troisième homme peut-il comprendre à quoi correspond « la même chose » alors qu'il n'a pas entendu le premier ? Peut-être que ce n'est pas le premier coup de fil qu'il reçoit, mais ce qui semble plus cohérent avec l'ensemble du film est que le regard de la caméra, celui du film, est celui du personnage-rêveur. Ce n'est donc pas la peine de répéter l'information, puisque nous l'avons déjà entendue par le premier homme.

Par conséquent, il semble que cette incohérence acoustique de la dernière sonnerie du téléphone correspond elle aussi au point de vue – point d'écoute subjectif du personnage principal, celle-là même qui arrive à l'aéroport dans la séquence suivante, émerveillée.

*i. Un silence non diégétique : Bande à part*

Dans la même séquence de *Bande à part* dont nous avons cité un extrait plus tôt, les personnages sont à table avant de danser. N'ayant rien à se dire, ils décident de faire « une minute de silence. » Un, deux, trois : tous les sons du film se coupent alors. On n'entend plus la rumeur du restaurant, plus les mouvements des personnages, simplement le souffle de la bande.

Cette minute de silence, qui ne dure en fait que 35 secondes, est interprétée littéralement par le réalisateur comme un silence cinématographique. Ce silence n'est pas diégétique. Les personnages entendent toujours le restaurant.

Il est difficile de faire rentrer cette expérience particulière dans les catégories que nous avons décrites jusqu'ici. Mais de manière moins extrême, le fait de ne pas entendre dans le film quelque chose que les personnages entendent dans la diégèse est un procédé courant dans le cinéma. Il est parfois d'intérêt dramatique, mais quasi systématique lorsque l'on compare ce que l'on entend dans le film à ce que l'on entend sur le lieu réel du tournage. D'une certaine façon, ne pas entendre de voitures dans un film se déroulant à Paris est un problème diégétique. Quoi qu'il en soit, il s'agit d'un mécanisme de dissociation entre la logique cinématographique et la logique diégétique.

Peut-être est-ce là le véritable « hors champ sonore » ? Comme lorsqu'on ne voit pas ce que les personnages voient, parfois, on n'entend pas ce qu'ils entendent.

j. *La diégèse ne préexiste pas au montage : Village, silenced*

*Village, silenced* (Deborah Stratman, 2012) est un court film dans lequel la réalisatrice reprend des images et des sons du film *The Silent Village* de Humphrey Jennings sorti en 1943. Dans ce court-métrage de 36 minutes, le réalisateur met en scène les habitants du village minier de Cwmgiedd au Pays de Galles reconstituant des événements s'étant déroulés dans un village similaire nommé Lidice en République Tchèque. Face à l'invasion nazie, les habitants organisent la résistance.

Le film de Deborah Stratman est la reconstruction d'une nouvelle séquence à partir de diverses images du film, prises surtout dans une séquence similaire au début du film, mais alimentée également par d'autres plans. Cette séquence est montée trois fois de suite : la première fois, elle est surtout accompagnée de sons du film original auxquels en ont été ajoutés d'autres ; la seconde, les sons sont surtout constitués d'alarmes ; la troisième est absolument silencieuse.

Afin de clarifier l'explication, nous décrirons d'abord la séquence originale du film de Humphrey Jennings, ensuite chaque itération du film de Deborah Stratman.



**Figure 9** *The Silent Village*

Dans le film original, accompagnée d'une musique de fanfare, une voiture surmontée d'un gros pavillon arrive dans le village, sous le regard des habitants qui continuent leurs activités. La voiture s'arrête alors et on entend un message en anglais, mais avec un accent allemand, annonçant que le Pays de Galles est à présent occupé par l'Allemagne et donnant des instructions concernant le nouvel ordre. Une fois cette partie du message terminée, la fanfare reprend et la voiture poursuit son chemin avant de s'arrêter de nouveau un peu plus loin dans le village et reprendre son annonce. On peut voir notamment à ce moment un villageois fermer sa fenêtre, comme s'il ne souhaitait rien entendre. La voiture repart enfin au loin, accompagnée d'un fondu au noir.

Dans le film de Deborah Stratman maintenant, la première itération de cette séquence est légèrement différente. Pas de fanfare cette fois, mais plutôt une ambiance de campagne, avec l'eau d'une rivière sous le pont où passe la voiture, et le moteur de cette même voiture, mixé différemment selon les plans. C'est déjà un montage de sons bien plus moderne que le film original, malgré la qualité technique identique. On entend ensuite un chien aboyer au passage de la voiture. Puis des oiseaux, un corbeau, tandis que la voiture poursuit sa route dans le village. Elle s'arrête alors, mais cette fois le message que l'on entend est différent : il s'agit d'un appel à la délation des assassins d'un général allemand. C'est un message que l'on entend bien plus tard dans le film

original de Jennings. Mais dans le film original, les habitants sont préoccupés par cette annonce. Ici, sur ces images, ils continuent leurs affaires, écoutant presque distraitement la voix qui vient de la voiture. Lorsque la voiture arrive devant l'école, on entend les enfants chanter depuis l'extérieur d'abord, puis à l'intérieur et s'arrêter à son passage en tournant la tête vers la fenêtre. A ce moment alors, le message change et reprend la deuxième partie du message original que l'on a entendu dans la première version. La voiture s'en va alors, au loin, mais coupe avant le fondu au noir. C'est la fin de la séquence et les images recommencent pour la deuxième version du film de Stratman.

Avant de poursuivre sur cet autre montage, nous pouvons déjà remarquer que les deux séquences analysées jusqu'ici ne possèdent que des sons diégétiques (à part peut-être la première fanfare où un doute existe). La deuxième version, malgré la qualité technique identique, est montée de manière bien plus moderne, avec des changements de plans sonores selon les plans visuels, des sons hors-champ traités fréquemment, de la réverbération sur les murs du village, etc. Ainsi, elles se ressemblent mais possèdent des différences à la fois sur les sons qui sont montés et sur la façon dont ils le sont.

Le deuxième montage de Stratman démarre avec des sons de coups de feu, au loin, et des bombes qui explosent. On entend un fort son de klaxon lorsqu'on se rapproche de la voiture ; il semble sortir du pavillon. Une alarme qui varie tonalement s'ajoute alors, mais comme si elle était plus loin. Il est difficile de dire si elle est diégétique, sortant peut-être de hauts parleurs dans le village, ou si elle n'est pas diégétique. Reviennent alors des sons de coups de feu et de bombes au loin, puis un nouveau klaxon, puissant, dont la source n'est pas claire. Apparaissent ensuite des sons stridents, s'ajoutant aux autres et formant une partition musicale violente. Les gros plans sur le pavillon semblent signifier que ces alarmes en proviennent. Les villageois la regardent, comme s'ils comprenaient ce qu'elle faisait ici, que cette voiture diffusant ces sons très forts était un spectacle habituel ; jusqu'à la fin de la séquence, où l'homme qui ferme sa fenêtre stoppe de son geste un son, puis une autre porte du village qui se ferme arrête totalement l'alarme, laissant place à une ambiance de campagne avec des oiseaux, un corbeau, tandis que la voiture s'éloigne, le moteur ronronnant.

Ce qui est frappant dans cette version, c'est d'abord la façon dont des sons qui semblent ne pas être diégétiques le deviennent par le phénomène de synchrèse que nous avons cité plus haut, particulièrement sur les gros plans du pavillon de la voiture. En voyant un

objet qui est censé diffuser du son, on associe les sons que l'on entend au même moment, en l'occurrence les alarmes, à cet objet. Cette partition musicale abstraite apparaît alors comme une agression sonore perpétrée par la voiture envers ce village. Mais les habitants ne semblent pas surpris par cette voiture ; leur absence de surprise, due dans la séquence originale au message allemand tout à fait compréhensible, laisse entendre cette fois que ce spectacle n'est pas nouveau. Tout un passé de passages de cette voiture dans ce village, diffusant ces mêmes sons, se crée alors.

Enfin, la dernière version de la séquence dans le film de Stratman est plus simple à décrire, puisque totalement muette. Ce sont les mêmes images, mais sans aucun son. Elles semblent cette fois laisser le spectateur libre d'imaginer ce que cette voiture diffuse, et ce que nous avons vu précédemment permet d'entrevoir l'immensité des possibilités.

Ce que ce film met en évidence est un résultat important : la diégèse ne préexiste pas au montage. En d'autres termes, il n'existe pas de diégèse *a priori* que l'on peut examiner pour savoir si un son que l'on souhaite monter y appartient ou non. C'est le montage lui-même qui définit ce qui est diégétique et ce qui ne l'est pas. Une même séquence, mais avec deux montages de sons différents, peut présenter deux diégèses totalement différentes ; c'est ce que nous avons vu ici.

## **8. Partie pratique**

Afin de mettre en application ce qui est décrit dans ce mémoire, je m'attèlerai à la fin des deux premières parties de ce mémoire à une analyse *a posteriori* de *Summertime*, le film que j'ai réalisé en tant que partie pratique de ce travail. Cette analyse nous permet de tenter de comprendre pourquoi nous avons fait certains choix, comment ils s'inscrivent dans un langage déjà existant, et ce que nous aurions pu faire autrement.

Je tâcherai donc ici de relever l'ensemble des sons non diégétiques que nous avons utilisé dans le montage de ce film. Ils sont classés dans l'ordre chronologique, avec leurs minutages de début, afin de simplifier le suivi.

Minutage de début	Son	Niveau diégétique	Types
00 :07	Musique <i>Greensleeves</i>	Non diégétique	Musique
00 :29	Voix de Benjamin, Luce et Pauline	Incohérence acoustique	Voix parlée / Son abstrait
01 :17	Nappe grave	Non diégétique	Son abstrait
02 :00	Luce tombe sur la table	Incohérence acoustique	Son abstrait
02 :13	Musique <i>Sèvres</i>	Non diégétique	Musique
03 :56	La porte grince	Incohérence acoustique	Son réaliste / Son abstrait
08 :40	Nappe grave	Non diégétique	Son abstrait
08 :48	Transmission radio	Non diégétique	Son réaliste / Son abstrait
09 :34	Nappe visage	Non diégétique	Son abstrait
10 :34	Voix Pauline	Incohérence acoustique	Voix parlée
11 :01	Nappe aigue clé	Non diégétique	Son abstrait
11 :40	Nappe aigue clé	Non diégétique	Son abstrait
12 :18	Musique de la mère	Non diégétique	Musique
13 :49	Nappe grave mallette et jeune homme	Non diégétique	Son abstrait
13 :51	Voix mère	Incohérence acoustique	Voix parlée / Son abstrait
13 :53	Passages de voitures	Incohérence acoustique	Son réaliste
14 :39	Ouverture de la porte, extérieur	Non diégétique	Son abstrait
14 :46	Fermeture porte	Incohérence acoustique	Son réaliste
14 :47	Nappes venteuses, toute la séquence	Non diégétique	Son abstrait
15 :45	Retournement	Non diégétique	Son abstrait
15 :57	Nappe grave	Non diégétique	Son abstrait
16 :33	Musique mallette	Non diégétique	Musique
16 :53	Clé	Incohérence acoustique	Son réaliste

<b>17 :12</b>	Ouverture mallette	Non diégétique	Son abstrait
<b>17 :14</b>	Mallette vide, arrivée Pauline	Non diégétique	Son abstrait
<b>17 :22</b>	Nappe venteuse	Non diégétique	Son abstrait
<b>18 :53</b>	Voix personnages sur la séquence	Incohérence acoustique	Voix parlée
<b>22 :10</b>	Extinction	Non diégétique	Son abstrait
<b>22 :14</b>	Nappe grave	Non diégétique	Son abstrait
<b>22 :16</b>	Transmission radio	Non diégétique	Son réaliste / Son abstrait
<b>22 :21</b>	Pas saxophoniste	Incohérence acoustique	Son réaliste
<b>24 :16</b>	Nappe venteuse	Non diégétique	Son abstrait
<b>26 :41</b>	« Unravel » - Björk	Non diégétique	Musique

L'ensemble des sons de la dernière séquence, à l'exception de la musique, ne sont pas présents dans ce tableau parce qu'ils correspondent au moment où le personnage principal se met à les entendre, ou au moins à entendre certains. La distinction entre diégétique et non diégétique devient floue et n'est pas déterminée.

## 9. Conclusion : artificialité et caractère diégétique

Les deux points importants que nous avons développés dans cette partie sont d'une part l'artificialité de la diégèse et d'autre part les différents critères qui permettent de déterminer le caractère diégétique d'un son.

La diégèse est une fabrication. Elle ne préexiste pas au montage. Un son n'appartient pas *a priori* à la diégèse ; cette appartenance dépend du contexte. En d'autres termes, un même son peut être diégétique dans un film et non diégétique dans un autre, voire même dans une autre séquence du même film.

Quant au caractère diégétique du son au sein d'un film, le choix n'est pas binaire. Il existe une succession de niveaux diégétiques, au sein desquels un son peut se situer clairement (comme une musique non diégétique par exemple), mais peut également se déplacer au cours du film (comme une musique diégétique devenant non diégétique, ou

le contraire), ou bien appartenir à plusieurs niveaux à la fois (comme une musique non diégétique tout à coup entendue par un personnage).

Ainsi, nous avons ajouté à la distinction classique entre diégétique et non diégétique un ensemble de niveaux intermédiaires : les sons à caractère non diégétique. Le caractère diégétique d'un son se représente mieux par un axe orienté plutôt que dans des cases fixes, et cette situation se résume par le schéma ci-dessous qui sera complété dans la suite de ce mémoire :



Dans la seconde partie de ce mémoire, nous étudions les fonctions narratives de ces sons non diégétiques.

## II. Fonctions narratives des sons non diégétiques

Nous l'avons vu, la diégèse est une fabrication. Par le montage, certains éléments de la bande sonore deviennent diégétiques, et d'autres non. Selon ce qui fait qu'un son n'appartient pas à l'histoire, on peut le qualifier de plus ou moins diégétique. Quoi qu'il en soit, il participe à la fabrication du récit, au processus de narration de l'histoire ; nous nous intéressons dans cette partie aux fonctions narratives des sons non diégétiques.

Nous fondons cette étude principalement sur la théorie de la narratologie développée par André Gaudreault et François Jost dans *Le Récit cinématographique*. Ce livre traite de l'ensemble des concepts d'analyse des moyens du récit, en reprenant les travaux de nombreux auteurs ayant fait avancer cette branche de la recherche, en premier lieu Gérard Genette, mais aussi entre autres Francis Vanoye et Christian Metz. De la même façon que Metz, qui l'exprime en 1968 dans les *Essais sur la signification au cinéma*, ce travail concerne « la sémiologie du film narratif » plutôt que « l'analyse structurale de la *narrativité* elle-même »<sup>6</sup>. En d'autres termes, il s'agit d'étudier la façon dont une histoire est racontée au cinéma, plutôt que les mécanismes internes à cette histoire.

Une difficulté de l'étude du récit cinématographique est que, comme le dit Jost, « l'image montre, mais ne dit pas ». Un plan contient une multitude d'énoncés possibles. Si l'on voit un personnage tomber, cela peut vouloir dire « il tombe », mais aussi, si l'on sait que le sol est glissant, « il glisse », ou bien « il fait une crise cardiaque » si le personnage est malade. Bref, il n'existe pas de traduction immédiate en mots du langage cinématographique. Malgré tout, cela n'empêche pas de mettre en évidence des mécanismes du récit filmique.

Le livre de Gaudreault et Jost présente toutefois certaines limites par rapport à notre sujet, celui des sons non diégétiques. Les auteurs ont tout à fait conscience que l'association d'images et de sons crée un nouveau langage et une attitude de perception nouvelle (le terme « audiovisualisation » est mentionné dès l'introduction), mais leur approche est selon nous incomplète :

- Les images et les sons sont surtout traités séparément, principalement lorsque cela concerne les questions de focalisation. Les plans revêtiraient en eux-mêmes la focalisation du récit, alors que nous sommes d'avis que cette

---

<sup>6</sup> *Le Récit cinématographique*, André Gaudreault, François Jost, p. 12

focalisation dépend fondamentalement de l'association entre ces plans et des sons.

- L'approche du rôle du son dans le récit cinématographique concerne principalement le langage. Les auteurs posent même la question : « est-ce que les bruits, et non plus les mots, peuvent être porteurs d'un récit ? On a vu [...] qu'ils contribuaient à sa construction, sans nul doute. Peut-on aller plus loin ? »<sup>7</sup>. Ils présentent ensuite un exemple allant dans ce sens, mais cette utilisation du son autre que verbal en tant qu'outil de narration est qualifiée de « relativement rare ».

Nous proposons dans cette partie de poursuivre cette étude en complétant leurs travaux sur l'aspect narratif des sons non diégétiques ; certains sont verbaux (c'est ce que nous avons appelé « la voix parlée ») mais d'autres ne le sont pas. L'idée que nous défendons, et l'exemple qu'ils utilisent s'y conforme, est que *l'utilisation d'un son non diégétique dans un film de fiction est avant tout un outil narratif*.

Malgré ces limites, ces travaux demeurent d'une grande richesse et nous commencerons par reprendre certaines notions qu'ils développent. Ensuite, nous verrons comment la fonction narrative du son non diégétique dépend d'une nouvelle notion que nous introduirons : la distance à l'histoire.

## 1. Le grand imagier

L'étude du récit cinématographique et des rapports entre l'énoncé et l'instance productrice amène rapidement à se poser une question simple : qui raconte le film ?

Selon la terminologie proposée par Albert Laffay dans *Logique du cinéma* en 1964, faisant de lui un précurseur dans la recherche narratologique, Gaudreault et Jost nomment la plus haute instance narratrice du film le « grand imagier ». Il faut comprendre ici un méga-narrateur qui engloberait tous les moyens de narration du cinéma, un narrateur fondamental, bref qui serait la source et la cause de la narration. Pour simplifier le discours, nous appellerons cette instance le réalisateur.

---

<sup>7</sup> *ibid.*, p. 30

Gaudreault propose de séparer la narration cinématographique en deux sous-instances : la *monstration* et la *narration*.

La monstration correspond au procédé même de la caméra, puis de la diffusion, celui de capter ce qui se passe afin de pouvoir le montrer de manière animée sur un écran. Il s'agit de photographier des photogrammes puis de les diffuser à la suite, créant par là même le plan.

La deuxième instance narrative, appelée tout simplement la narration, consiste à agencer ces plans de manière à créer des liens entre eux et à raconter une histoire : c'est le montage.

En nous inspirant d'un schéma de Gaudreault, nous pouvons résumer la situation ainsi :

Méga-narrateur (grand imagier / réalisateur)	Monstrateur (caméra)
	Narrateur (montage)

Cette classification n'envisage pas le son dans sa construction. Mais cela ne la remet pas totalement en cause, puisque nous proposons d'y classer les sons de cette façon :

- Si le son est diégétique et commence et finit en même temps qu'un plan, c'est-à-dire qu'il ne déborde pas d'un plan à un autre, il appartient à la monstration. Cela est indépendant du fait qu'il s'agisse de son direct ou pas : l'important est l'illusion d'unité entre les caractères visuels et sonores du plan.
- Si le son est diégétique mais chevauche les plans, ou s'il n'est pas diégétique, il appartient à la narration. C'est par une opération de montage qu'il prend sa place dans le récit.

Cela se résume ainsi :

Son diégétique		Son non diégétique
Ne chevauche pas de plan	Chevauche des plans	Narration
Monstration	Narration	

Les sons non diégétiques appartiennent donc à l'instance de la narration. Cela n'est pas surprenant : puisqu'ils n'appartiennent pas à l'histoire, ils ne peuvent pas y être captés.

Qu'en est-il alors des sons à caractère non diégétique ? Cela pose problème car ils peuvent à la fois appartenir à un plan, et en même temps être dissociés de la diégèse par un ou plusieurs de leurs caractères. Nous ne pouvons apporter de réponse complète, mais nous pouvons émettre l'hypothèse qu'ils sont partagés par les deux instances, qu'ils relèvent à la fois de la monstration et de la narration.

## 2. Le hors-champ diégétique

Gaudreault et Jost mentionnent un « espace hors-champ, et qui est directement sous la coupe du grand imagier », en faisant référence à un mélange de ce que Laurent Jullier appelle les mondes intérieurs et les fosses imaginaires. Il s'agit en fait du lieu hypothétique d'où proviennent les sons non diégétiques et à caractère non diégétique.

C'est un espace envisagé comme un hors-champ diégétique, celui d'où proviennent les sons qui n'appartiennent pas à l'histoire. En pratique, nous savons bien qu'un tel espace n'existe pas et que les sons non diégétiques proviennent de diverses sources enregistrées à des endroits et des moments différents. Dans l'étude narratologique cependant, cet espace, défini d'ailleurs uniquement par sa différence avec l'espace diégétique, prend son sens car il permet l'utilisation de sons dans la narration (nous l'avons vu ci-dessus, le montage) qui ne relèvent pas de la monstration.

Cela n'est pas problématique. Il suffit simplement de considérer le monstreur comme une instance plus large qui ne se contente pas de filmer des plans, mais qui enregistre également d'autres choses ailleurs, à d'autres moments. Le plan n'apparaît plus comme la seule fabrication du monstreur. Il faudrait plutôt parler de *rush*, dans le sens pratique que nous utilisons dans la fabrication des films, c'est-à-dire d'une part les plans, avec le son synchronisé, mais également les sons seuls, les bruitages, et plus généralement l'ensemble de la matière sonore enregistrée indépendamment de l'image.

### 3. La focalisation chez Genette

« Pour éviter ce que les termes de *vision*, de *champ* et de *point de vue* ont de trop spécifiquement visuel », Genette propose le terme de *focalisation*. Ce terme a été choisi pour signifier que son objectif est de déterminer le *foyer* du récit (*focus* en anglais)<sup>8</sup>. Ainsi, Genette détermine trois focalisations différentes du récit :

- Récit *non focalisé* ou à *focalisation zéro*, le cas d'un narrateur omniscient ;
- Récit à *focalisation interne*, lorsqu'il se met à la place d'un personnage, que l'on sait ce qu'il sait ;
- Récit à *focalisation externe*, lorsque l'on n'a pas accès au savoir des personnages.

Jost précise ces notions en y ajoutant celles d'*ocularisation* et d'*auricularisation*, désignant respectivement la relation entre ce que l'on voit et ce que le personnage est censé voir et la relation entre ce que l'on entend et ce que le personnage est censé entendre. La focalisation se limite quant à elle à la relation entre le savoir du narrateur et celui du personnage. Nous n'entrerons pas dans le détail de ces notions, car la tripartition de Genette nous sera suffisante pour examiner le cas des sons non diégétiques.

### 4. Configurations discursives du point de vue

Plutôt que d'étudier l'auricularisation des sons non diégétiques, il s'agit d'examiner les rapports entre énonciateur, énonciataire (l'instance à qui s'adresse l'énoncé) et énoncé. Nous utilisons pour cela la terminologie proposée par le chercheur italien Francesco Casetti<sup>9</sup>, qui identifie dans un texte (mais cela vaut aussi pour le discours filmique) l'énonciateur au *Je*, l'énonciataire au *Tu* et l'énoncé au *Il*.

Dans le cadre du cinéma, le *Tu*, c'est-à-dire la personne à laquelle s'adresse le plan, est toujours le spectateur. Le *Je* peut être variable, nous allons le voir. Le *Il* correspond à ce qui est montré dans le plan.

---

<sup>8</sup> *ibid.*, p. 128

<sup>9</sup> *ibid.*, p. 58-59

Sont déterminées ainsi quatre « configurations discursives du point de vue », selon les types de plans du cinéma :

- La configuration objective : les plans qui n'adoptent pas de point de vue particulier, c'est-à-dire la plupart des plans, dont le sujet principal est ce qui est montré ; le *Il* prime sur le *Je* et le *Tu*.
- La configuration de l'adresse : les plans s'apparentant au regard caméra, lorsque le personnage à l'écran semble contrôler la caméra qui le montre et s'adresse au spectateur ; le *Je* et le *Il* se confondent et s'adressent au *Tu*.
- La configuration subjective : l'exemple du plan subjectif ; le *Tu* et le *Il* ne font plus qu'un et voient ce que le *Je* donne à voir.
- La configuration objective irréelle : « lorsque la caméra manifeste de façon ostensible son omnipotence », c'est-à-dire que le plan dit en lui-même qu'il est produit par un narrateur (un monstateur), mais que ce n'est pas un personnage de l'histoire ; cette fois c'est le *Je* qui prime et qui signifie au *Tu* qu'il choisit de montrer ce qu'il veut.

Ces concepts assez abstraits sous cette forme correspondent à une idée simple : où (dans quoi) est contenu le *sens* du plan? Autrement dit, pourquoi est-ce que ce plan existe, que donne-t-il à voir, que doit-on comprendre ?

Dans le cas des sons non diégétiques, que nous considérons de la même façon comme un énoncé, l'analogie suivante peut être établie :

- Le *Je* peut correspondre au réalisateur (le grand imagier) ou à un personnage ;
- Le *Il* correspond au son en lui-même ;
- Le *Tu* au spectateur.

Plutôt que de les classer dans des configurations figées, il s'agit d'étudier à quel point chacun de ces éléments (*Je*, *Tu*, *Il*) est important dans l'utilisation de ce son, et plus précisément leurs rapports d'importance. Nous allons le voir un peu plus tard, la fonction narrative d'un son non diégétique est intimement liée à ces rapports.

## 5. La double temporalité

Le dernier concept essentiel que nous tirons des travaux de Gaudreault et Jost est celui de la double temporalité du récit : d'une part celle de l'histoire et d'autre part celle du récit, en tant que l'acte de raconter l'histoire<sup>10</sup>.

Dans une histoire, un événement se produit à un certain moment, pendant un certain temps, et un certain nombre de fois. Par exemple :

- « Daniel se lève tous les matins à sept heures » : quand ? = à sept heures, combien de temps ? = le temps de se lever, combien de fois ? = tous les matins ;
- « Mais aujourd'hui, Daniel décide de rester au lit jusqu'à huit heures » : quand ? = aujourd'hui, combien de temps ? = de sept à huit heures, combien de fois ? = une seule fois.

Le récit possède sa temporalité propre, et n'est pas obligé de suivre strictement celle de l'histoire. La plupart du temps, elle en est même très différente.

Dans l'exemple que nous avons pris ci-dessus, un récit écrit, le « quand » correspond au moment où on le lit, le « combien de temps » à la durée de la lecture, et le « combien de fois » une seule à chaque fois. Au cinéma, les mêmes comparaisons peuvent être réalisées.

Ainsi, les temporalités de l'histoire et du récit peuvent être comparées à trois niveaux :

- L'ordre dans lequel les événements se succèdent ;
- La durée qu'ils ont dans l'histoire et celle avec laquelle ils sont racontés ;
- La fréquence, le nombre de fois où ils se produisent dans l'histoire et le nombre de fois où ils sont évoqués dans le récit.

Cette double temporalité nous concerne particulièrement dans le cas des sons non diégétiques, parce que n'appartenant pas à l'histoire, ils n'entrent donc pas dans le cadre de sa temporalité. Etant étrangers à l'histoire, les sons non diégétiques ne sont présents que dans la temporalité du récit.

---

<sup>10</sup> *ibid.*, p. 104

Il est cependant tentant de vouloir les placer dans la temporalité de l'histoire. En effet, étant montés dans le film en même temps que des événements diégétiques, il semble logique d'imaginer que d'une certaine façon, ils existent dans l'histoire au même moment. Cela semble particulièrement vrai dans le cas de sons non diégétiques ponctuels, montés de manière synchrone à des événements de l'histoire. Pour reprendre un exemple de la première partie de ce mémoire, lorsque la dernière sonnerie du téléphone de *Mulholland Drive* résonne beaucoup plus que les autres, il semble que ce son à caractère non diégétique appartient tout de même d'une certaine façon à l'histoire, ou qu'en tout cas on peut le placer dans la temporalité de l'histoire.

Mais il ne faut pas se tromper. Les sons non diégétiques sont qualifiés ainsi précisément parce qu'ils n'appartiennent pas à l'histoire, mais uniquement au récit. C'est une caractéristique très particulière des sons non diégétiques, qui n'est pas partagée par beaucoup d'autres éléments du récit. On peut penser au générique, aux sous-titres, aux cartons par exemple. A chaque fois cependant, l'énonciateur est clairement extérieur à l'histoire, et ce n'est pas toujours le cas des sons non diégétiques.

Nous avons affaire à une contamination de la temporalité de l'histoire par celle du récit. Cette aspiration particulière qu'a le son non diégétique en tant qu'élément du récit à devenir un élément de l'histoire, nous proposons de l'appeler *distance à l'histoire*.

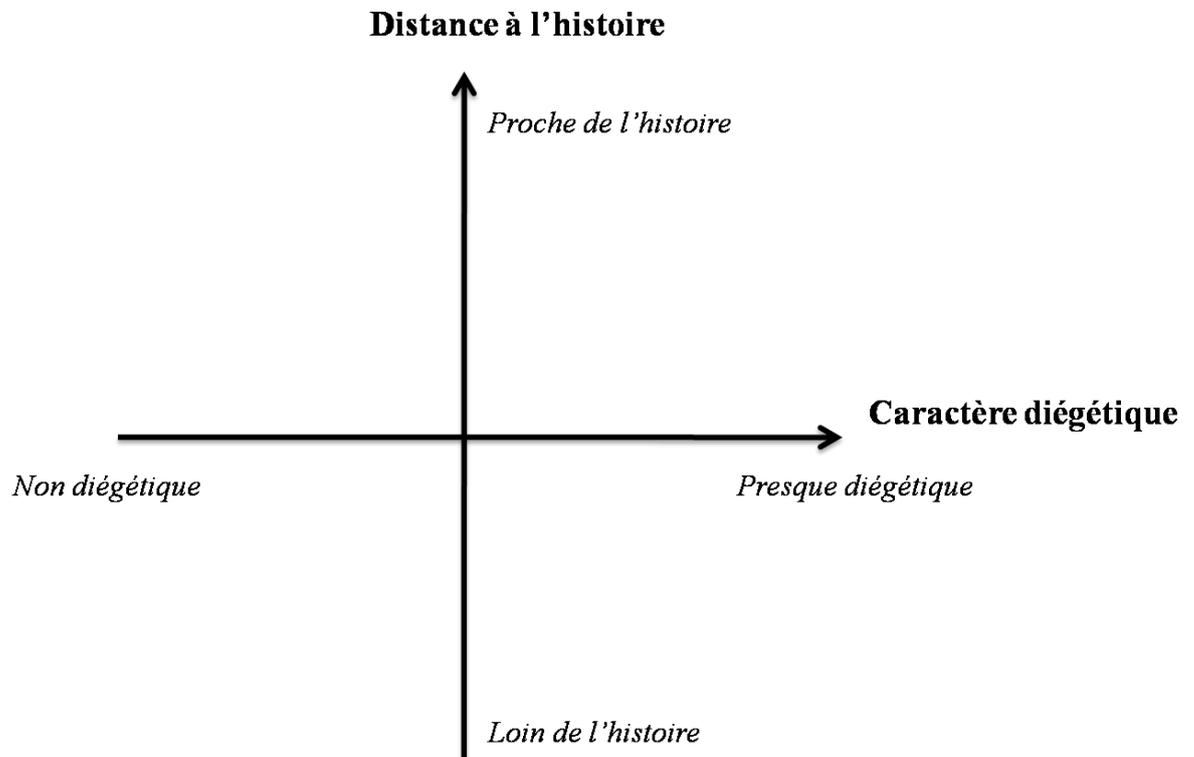
## 6. La distance à l'histoire : définition

La distance à l'histoire est l'expression que nous proposons pour qualifier le rapport d'un son non diégétique à l'histoire. On dira d'un son qu'il est *proche* de l'histoire lorsque la diégèse semble ne pas pouvoir exister sans ce son. *A contrario*, un son est *loin* de l'histoire lorsque la diégèse semble indépendante de lui, qu'il n'y a aucun doute sur le fait qu'elle existe sans lui.

Cette distance à l'histoire est absolument distincte du caractère diégétique du son. Un son dont le caractère non diégétique est évident peut tout à fait être très proche de l'histoire. C'est le cas par exemple de la musique de *Star Wars* dont nous avons pris l'exemple plus tôt. Lorsque le personnage de Dark Vador apparaît à l'écran, il est intimement lié à son thème musical ; cette musique fait partie du personnage autant que

son costume. Pourtant, il n'y a aucun doute sur le fait qu'elle est absolument non diégétique.

Nous pouvons donc compléter le schéma que nous avons établi dans la première partie de ce mémoire avec un nouvel axe :



Cette distance à l'histoire n'est pas une donnée strictement mesurable ; elle dépend à la fois de la perception du spectateur, de son rapport au film, et de la place que le son occupe au sein du film dans sa globalité. Ainsi, certains spectateurs peuvent dire après le visionnage d'un film qu'il y avait beaucoup de musique, tandis que d'autres ne s'en sont pas rendus compte.

La distance d'un son à l'histoire dépend à la fois de critères spécifiques au type de son non diégétique auquel il appartient, ainsi qu'à des critères généraux partagés par tous. Nous allons les présenter successivement.

## 7. La distance à l'histoire : critères spécifiques

La distance à l'histoire dépend donc en premier lieu de plusieurs critères spécifiques à chaque type de son non diégétique.

### a. *La musique*

Laurent Jullier distingue trois grandes familles de styles de musiques au cinéma, qui « délimitent moins des époques que des façons de concevoir le cinéma et son impact sur le spectateur »<sup>11</sup>. Ces trois familles sont la musique *classique*, la musique *moderne* et la musique *postmoderne*. Cette répartition rejoint notre concept de distance à l'histoire car cette distinction se fonde surtout sur le rapport que la musique crée entre le spectateur et le film.

De la musique classique, Jullier dit qu'elle « favorise l'homéostasie, c'est-à-dire l'équilibre entre émotion et réflexion »<sup>12</sup>. Dans ce cas, la musique est au service de la narration, afin que l'histoire soit comprise et ressentie par le spectateur. La musique classique est *proche de l'histoire*.

Au contraire, la musique moderne place le spectateur à distance. Elle rend évident le fait qu'une histoire est en train d'être racontée. Selon la classification de Casetti, l'énonciateur *Je*, c'est-à-dire le réalisateur, est signifié fortement par ce type de musique. La musique moderne est *loin de l'histoire*.

Enfin, la musique postmoderne occupe une position trouble. Jullier dit d'elle qu'elle est « au service du vertige et du *bain de sons* »<sup>13</sup> ; autrement dit, elle joue presque un rôle d'ambiance, comme la créatrice d'un univers singulier propre au film. Elle se rapproche de ce que nous avons appelé les sons abstraits. Selon d'autres critères généraux que nous distinguons plus tard, elle peut être proche ou loin de l'histoire.

### b. *La voix parlée*

La distance à l'histoire de la voix parlée dépend surtout de la focalisation qu'elle adopte, et de l'importance qu'y occupent le *Je* et le *Tu*, l'énonciateur et l'énonciataire, par rapport au *Il*, l'énoncé. La tendance générale est que plus l'énoncé sera prédominant, plus la voix parlée sera proche de l'histoire ; au contraire, plus l'énonciateur et l'énonciataire seront importants, plus la voix parlée en sera éloignée.

---

<sup>11</sup> *Le son au cinéma*, Laurent Jullier, p. 48

<sup>12</sup> *ibid.*, p. 48

<sup>13</sup> *ibid.*, p. 49

c. *Les sons réalistes*

Le critère le plus important pour déterminer la distance à l'histoire d'un son réaliste est la probabilité qu'il appartienne à la diégèse, leur rapport culturel. Si dans un autre contexte le son aurait pu être diégétique, il aura tendance à être plus proche de l'histoire que s'il y est étranger.

Nous le verrons également avec les sons abstraits, le rapport de synchronisme entre un événement diégétique et un son réaliste non diégétique rapproche le son de l'histoire. C'est le phénomène de synchrèse que nous avons cité plus tôt : le son, même s'il n'est pas diégétique, semble être lié indéfectiblement à l'image.

d. *Les sons abstraits*

La distance à l'histoire des sons abstraits, encore plus fortement que pour les sons réalistes, dépend essentiellement de la synchrèse. Plus précisément, un son abstrait sera d'autant plus proche de l'histoire qu'il essaye de se faire passer pour un son diégétique.

## 8. La distance à l'histoire : critères généraux

Les critères spécifiques que nous avons exposés ne sont pas suffisants pour décrire pleinement le rapport à l'histoire des sons non diégétiques. Certains critères plus généraux les complètent, voire prévalent dans certains cas.

a. *Le mixage*

La distance à l'histoire d'un son non diégétique est un moyen de déterminer à quel point il appartient à cette histoire. Il y a là une idée d'incorporation que l'on retrouve dans le mixage.

Le rapport entre le son non diégétique et l'histoire est très lié au rapport entre ce son non diégétique et les autres sons diégétiques qui sont présents dans la séquence. Ainsi, si le son non diégétique est mixé de telle façon qu'il se confonde avec les sons

diégétiques, il aura tendance à être plus près de l'histoire que s'il en est distinct dans le mélange.

C'est une question de rapport de volumes avant tout. La façon la plus immédiate de faire ressortir un son dans le mixage est de le mettre plus fort que les autres. Le son non diégétique aura donc tendance à être d'autant plus distant de l'histoire qu'il sera plus fort que les autres sons de la séquence.

C'est également une question de spatialisation. L'exemple le plus évident est encore une fois celui de la musique. Si elle est spatialisée en stéréo, dans les enceintes latérales, elle sera d'autant plus décorrélée des sons diégétiques, dont l'essentiel du message provient de l'enceinte centrale.

#### *b. La culture*

L'utilisation des sons non diégétiques dans le cinéma de fiction est telle que certains types de sons sont rapidement associés par le spectateur initié à un genre, un style. Parfois même, c'est l'utilisation d'un certain son qui signifie la première qu'un film appartient à un genre.

Cette connaissance préalable du spectateur induit des attentes, qui peuvent être satisfaites ou évitées par le film. Si le son non diégétique est habituellement associé au genre du film, il aura tendance à se rapprocher de l'histoire ; si par contre il appartient à un autre genre, créant par son utilisation un décalage avec le genre du film, il aura tendance à s'éloigner de l'histoire.

#### *c. L'expérience personnelle du spectateur*

Ce critère se rapproche du précédent, mais concerne cette fois le rapport du spectateur au son non diégétique, plutôt qu'au rapport entre le son et le film.

Si le spectateur connaît déjà le son non diégétique utilisé, que ce soit une musique, une voix ou un son réaliste, il aura selon le contexte plus ou moins tendance à l'identifier immédiatement comme non diégétique. Par exemple, entendre une musique que l'on connaît déjà aura tendance à l'éloigner de l'histoire ; de la même façon, l'anachronisme

d'un coup de feu ou d'une voiture dans un film médiéval l'éloigne également de l'histoire. Ce critère concerne moins les sons abstraits, pour lesquels c'est plutôt l'effet inverse qui s'opère : entendre un son connu amènera plutôt le spectateur à ne pas l'analyser et le rapprochera par là même de l'histoire.

## 9. Fonctions narratives

Les rapports entre énonciateur (*Je*), énonciataire (*Tu*), et énoncé (*Il*), que nous avons décrits dans le cas particulier de la voix parlée, sont généralisables à l'ensemble de ces critères. L'énonciateur sera d'autant plus visible que le son non diégétique sera distant de l'histoire, et donc s'affirmera en tant qu'élément du récit. Ce qu'il faut examiner, c'est ce qui « sort » le spectateur de l'histoire.

Il ne faut pas y voir de jugement de valeur. Un son proche de l'histoire n'est pas « mieux » qu'un son loin de l'histoire. Leur différence réside dans la fonction qu'ils occupent dans le récit.

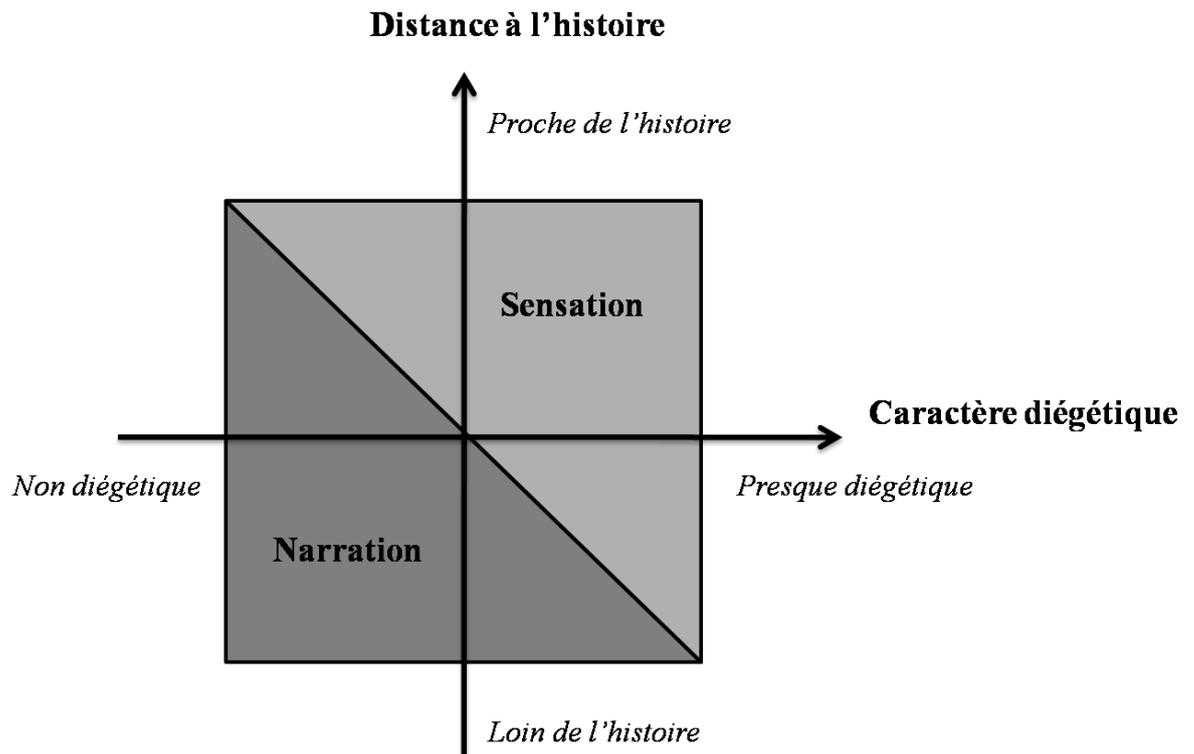
La tendance générale est la suivante :

- Un son non diégétique loin de l'histoire favorise *la narration*. Dans ce cas, le *Je* et le *Tu* prédominent. Le son rappelle au spectateur qu'on est en train de lui raconter une histoire, il le lui en fait prendre conscience.
- Un son non diégétique proche de l'histoire favorise *la sensation*. Ici, c'est le *Il* qui prédomine. Le son n'est pas identifié immédiatement par le spectateur comme extérieur à l'histoire. Il est pris pour lui-même, dans une attitude non analytique.

Cette tendance est pondérée par le caractère non diégétique du son, « à quel point » il n'est pas diégétique. Ainsi, en mêlant les notions de caractère diégétique et de distance à l'histoire, nous pouvons mettre en évidence quatre cas différents :

Caractère diégétique	Distance à l'histoire	Fonction
Non diégétique	Loin	Narration
Non diégétique	Proche	Surtout sensation, et narration
Presque diégétique	Loin	Surtout narration, et sensation
Presque diégétique	Proche	Sensation

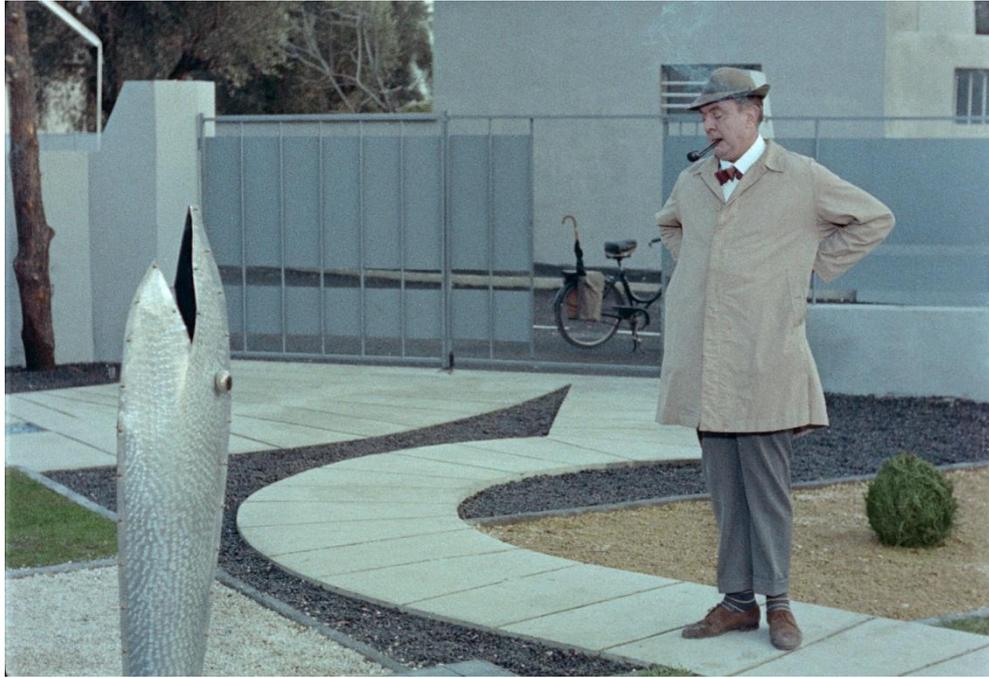
Cela nous permet de compléter le schéma :



## 10. Quelques exemples

### a. *Mon Oncle*

Dans *Mon Oncle* (Jacques Tati, 1958), l'absence de son direct au profit de bruitages qui miment les actions des personnages illustre bien le concept de distance à l'histoire.



**Figure 10** *Mon Oncle*

Prenons l'exemple précis des bruits de pas des personnages. Ce que nous avons développé dans la première partie de ce mémoire permet de qualifier ces sons comme étant presque diégétiques. Ils appartiennent à l'univers de l'histoire, même si leur aspect de fabrication non dissimulé permet de remettre en doute leur caractère diégétique. Ils appartiennent à la catégorie des sons réalistes.

Ils sont cependant loin de l'histoire. En effet, la façon dont ils sont mixés, le fait qu'ils ne soient pas intégrés à un bain de sons diégétiques, ne permet pas de doute sur le fait qu'ils ont été ajoutés au film par des opérations de montage. Le réalisateur en tant qu'énonciateur marque ainsi sa présence, et signifie en permanence au spectateur qu'il regarde un film. C'est la narration qui est privilégiée dans ce cas, dans le sens où entendre quelqu'un marcher signifie presque directement « il marche ».

*b. Oncle Boonmee, celui qui se souvient de ses vies antérieures*

*Oncle Boonmee, celui qui se souvient de ses vies antérieures* (Apichatpong Weerasethakul, 2010) est séparé en plusieurs parties, à chaque fois de la durée d'une bobine, présentant chacune une esthétique différente. La cinquième des six bobines se déroule en grande partie dans une grotte. Tout du long, une nappe grave et vrombissante emplît l'espace.



Figure 11 *Uncle Boonmee, celui qui se souvient de ses vies antérieures*

Ce son semble rapidement clairement non diégétique. Pourtant, son omniprésence et son importance tout au long de la bobine, et le lien indéfectible qui semble se créer entre ce décor et ce son le rendent très proche de l'histoire. S'il se posait la question, le spectateur remarquerait très vite que ce son n'est pas diégétique, mais le film n'appelle pas le spectateur à se la poser. L'énonciateur est très peu présent par l'usage de ce son, et c'est cette fois la sensation qui est privilégiée.

c. *Annie Hall*

*Annie Hall* (Woody Allen, 1977) présente tout au long du film une voix off du personnage principal, interprété par Woody Allen lui-même, s'adressant au spectateur. Ce cas de voix off pose question dans le cadre des sons non diégétiques, car il crée une communication entre le monde diégétique et le monde de la production matérielle, celui auquel les spectateurs appartiennent.

Cette voix off n'est pas diégétique, elle n'appartient pas à l'univers des scènes qui se déroulent dans le film, mais elle n'est pas non plus complètement non diégétique, car la personne qui la dit est le personnage principal du film, appartenant lui à la diégèse. Le fait qu'il s'adresse au spectateur signifie très fortement à ce dernier que quelque chose échappe à la logique diégétique.



**Figure 12** *Annie Hall*

C'est un exemple qui se rapproche de l'aparté au théâtre. Le personnage interrompt l'action diégétique qu'il est en train de mener et s'adresse immédiatement au spectateur, dans l'espace de la scène. C'est donc un son à caractère non diégétique qui est loin de l'histoire, qui favorise la narration, qui fait prendre conscience au spectateur qu'on lui raconte une histoire. Dans ce cas précis, cela est évocateur de l'expérience préalable de comédien de *stand-up* de Woody Allen.

## **11. Partie pratique**

De la même façon que dans la première partie, nous reprenons le même relevé des sons non diégétiques de la partie pratique de ce mémoire afin d'en examiner leur distance à l'histoire.

<b>Minutage de début</b>	<b>Son</b>	<b>Distance à l'histoire</b>
00 :07	Musique <i>Greensleeves</i>	Loin
00 :29	Voix de Benjamin, Luce et Pauline	Plutôt proche
01 :17	Nappe grave	Plutôt proche
02 :00	Luce tombe sur la table	Proche
02 :13	Musique <i>Sèvres</i>	Loin
03 :56	La porte grince	Proche
08 :40	Nappe grave	Proche
08 :48	Transmission radio	Proche
09 :34	Nappe visage	Proche
10 :34	Voix Pauline	Proche
11 :01	Nappe aigue clé	Plutôt proche
11 :40	Nappe aigue clé	Plutôt proche
12 :18	Musique de la mère	Plutôt loin
13 :49	Nappe grave mallette et jeune homme	Plutôt loin
13 :51	Voix mère	Plutôt loin
13 :53	Passages de voitures	Proche
14 :39	Ouverture de la porte, extérieur	Proche
14 :46	Fermeture porte	Proche
14 :47	Nappes venteuses, toute la séquence	Plutôt proche
15 :45	Retournement	Plutôt loin
15 :57	Nappe grave	Plutôt proche
16 :33	Musique mallette	Plutôt loin
16 :53	Clé	Proche
17 :12	Ouverture mallette	Plutôt loin
17 :14	Mallette vide, arrivée Pauline	Plutôt loin
17 :22	Nappe venteuse	Proche
18 :53	Voix personnages sur la séquence	Proche
22 :10	Extinction	Proche
22 :14	Nappe grave	Proche
22 :16	Transmission radio	Proche
22 :21	Pas saxophoniste	Proche
24 :16	Nappe venteuse	Plutôt proche
26 :41	« Unravel » - Björk	Loin

## 12. Conclusion : distance à l'histoire

Nous avons décrit dans cette partie les fonctions narratives des sons non diégétiques par l'intermédiaire d'une notion que nous avons introduite appelée la distance à l'histoire. Elle rend compte du lien entre un son non diégétique et la diégèse, et implique une réception différente de ce son par le spectateur, et donc un rapport au film différent.

Ce rapport au film se caractérise d'un côté par la prédominance de la narration lorsque le son est loin de l'histoire. Ces cas représentent les sons qui font prendre conscience au spectateur qu'une histoire est en train de lui être racontée. Il s'agit d'une certaine façon d'inciter le spectateur à comprendre quelque chose, comme pourrait le faire un mot de liaison ou une pause dans un discours oral.

De l'autre côté, lorsque le son est proche de l'histoire, la sensation du spectateur est favorisée par rapport à la narration. Dans ce cas, tout s'opère comme si le spectateur ne remarquait pas le son, ou plus précisément ne remarquait pas son caractère non diégétique. L'illusion de réalisme du film n'est pas perturbée par un son non diégétique proche de l'histoire, qui ne s'affirme pas immédiatement en tant qu'élément étranger à l'histoire.

Pondérée par le caractère diégétique du son, cette distance à l'histoire est également une donnée continue qui dépasse une distinction binaire. Elle peut être mouvante au sein d'une séquence, d'un film, et dépend comme le caractère diégétique du son de la perception qu'en a le spectateur.

A cette étape, nous avons étudié les sons non diégétiques uniquement par le prisme de leur nature, de la description de leur source, et indépendamment de leur esthétique. En comparant à la musique, c'est comme si nous étions capables de dresser la « partition » du film. L'objet de la dernière partie de ce mémoire est de mesurer l'effet de l'interprétation de cette partition, de la réalisation pratique des sons non diégétiques, sur la perception du film par le spectateur.

### III. Esthétique des sons non diégétiques

La dernière partie de ce mémoire est une étude plus pratique que les précédentes. Nous avons jusqu'ici considéré les sons non diégétiques comme des éléments du récit et examiné leurs fonctions narratives. Nous nous plaçons dans le cadre de l'analyse. En utilisant maintenant cette approche analytique dans un but pratique, c'est-à-dire celui de choisir et de monter des sons dans un film pour raconter une histoire, une difficulté majeure apparaît : quel son choisir ?

Afin d'éclairer le sens de cette question, prenons un exemple. Nous avons choisi de monter un quatuor à cordes sur une scène. Nous savons comment la mixer pour l'intégrer aux sons diégétiques de la séquence, mais la musique n'est pas enregistrée pour l'instant. Comment choisir la façon de l'enregistrer et les éventuels traitements à appliquer ? Est-ce que ces choix auront une importance sur la façon dont le film sera perçu par le spectateur ?

La question est très large et est commune à tous les arts. L'œuvre terminée, il est difficile d'en séparer les éléments qui la constituent pour expliquer les causes des émotions, des réflexions, bref des effets qu'elle peut susciter sur la personne qui la reçoit. De la même façon qu'un peintre peut se demander quelle nuance de bleu utiliser, la fabrication d'un film entraîne un très grand nombre de micro-choix qui changent l'objet final.

Par rapport aux autres éléments de la bande sonore d'un film, les sons non diégétiques présentent une approche différente car ils ne sont pas liés à des événements de l'image ; plus précisément, ils ne sont pas nécessairement liés à leur cause. Ils sont par nature inattendus par le spectateur, et permettent donc une liberté supplémentaire dans leur conception. Puisqu'avec les sons non diégétiques « tout est possible », c'est un terrain d'expérimentation privilégié qui justifie leur étude séparée.

Cette partie consistera surtout en une expérience basée sur une séquence de *Mulholland Drive* (David Lynch, 2001) que nous avons déjà étudiée plus tôt dans ce mémoire<sup>14</sup>. Avant d'en expliquer le protocole et l'objectif, nous présenterons quelques concepts théoriques qui la justifient.

---

<sup>14</sup> cf I. 7. f)

## 1. Pierre Schaeffer et les trois écoutes

Dans le *Traité des Objets Musicaux*, Pierre Schaeffer tente en 1966 de généraliser la musique à tous les sons et de créer un nouveau solfège (*Solfège de l'Objet Sonore*). Pour cela, il fait table rase de notions qu'il déconstruit, comme celle du timbre d'un instrument de musique, et propose une nouvelle conception de la musique fondée sur ce qu'il nomme *l'écoute réduite* et *l'objet sonore*<sup>15</sup>. Ces deux concepts sont absolument indissociables dans les travaux de Schaeffer, mais nous ne détaillerons ici que l'écoute réduite car son approche est au centre de l'expérience que nous présentons.

Schaeffer définit d'abord les deux écoutes classiques du son : l'écoute causale et l'écoute sémantique<sup>16</sup>. L'écoute causale a pour objectif de déterminer par les caractéristiques du son la source qui l'a émis. L'écoute sémantique s'intéresse quant à elle au message transporté par le son (Chion précise « un signal sonore codé »), dont l'exemple le plus fréquent est le langage.

Ces deux écoutes peuvent s'opérer en même temps. On peut écouter quelqu'un parler et comprendre ce qu'il dit (écoute sémantique), tout en remarquant par exemple que c'est un homme (écoute causale).

Sur ce dernier exemple, Schaeffer précise qu'on entend également que la voix de l'homme est *grave, rauque*. Cela l'amène alors à définir l'écoute réduite, qui s'intéresse aux caractéristiques du son en lui-même, indépendamment de la cause qui l'a émis et du message qu'il transmet. Cette écoute s'opère en fait naturellement en permanence, afin de donner des indications aux deux autres écoutes, mais la démarche de Schaeffer est de la poser en soi, afin de décrire ce qu'il définit alors comme l'objet sonore ; il s'agit de l'objet de l'écoute réduite : « tout phénomène 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<sup>17</sup> ».

Ces trois écoutes ont inspiré notre démarche de typologie des sons non diégétiques. En ce sens, nous avons considéré la musique comme privilégiant l'écoute sémantique (langage musical), la voix parlée privilégiant l'écoute sémantique également (langage

---

<sup>15</sup> *Le Son*, Michel Chion, p. 186

<sup>16</sup> *ibid.*, p. 187

<sup>17</sup> *Guide des objets sonores, Pierre Schaeffer et la Recherche Musicale*, Michel Chion, p.34, cité par Michel Chion lui-même dans *Le Son*, p. 189

verbal), les sons réalistes privilégiant l'écoute causale, et enfin les sons abstraits privilégiant l'écoute réduite.

Il est important de noter que l'écoute réduite, même si elle s'opère toujours en partie, ne peut être considérée seule que dans le cas de l'écoute de sons enregistrés. Afin de déterminer les caractéristiques d'un son, Schaeffer explique qu'il est requis de pouvoir l'écouter plusieurs fois, en boucle, afin qu'il se détache petit à petit de sa cause et du sens qu'il transporte, qu'il se vide du contenu des autres écoutes pour n'être plus qu'un objet sonore.

Dans le cas de notre étude, nous devons ici nous éloigner de la démarche de Schaeffer. Lorsqu'un spectateur regarde un film, les sons ne sont entendus qu'une seule fois, au moment où ils sont montés. Il est donc vain de prétendre pouvoir déterminer à la manière de Schaeffer les caractéristiques des sons d'un film. Cependant, les mécanismes de l'écoute réduite s'opèrent toujours lorsque nous percevons un son, et c'est précisément cette partie de l'écoute que nous étudions ici.

## **2. De l'importance de l'écoute réduite au cinéma**

L'écoute réduite a pour intérêt principal de considérer le son en lui-même et donne des informations aux deux autres écoutes. La question de son importance au cinéma est essentielle dans la façon d'aborder la fabrication d'un film.

Si l'on considère que l'écoute réduite n'a pas d'importance au cinéma, et qu'elle n'est utilisée qu'au service de la narration, pour comprendre ce qu'il se passe, alors dans ce cas il suffit qu'un son nous évoque sa source suffisamment bien pour qu'il soit convenable (écoute causale) ou il suffit qu'une voix soit intelligible pour qu'elle assure son rôle (écoute sémantique). Cela voudrait dire, dans ce cas toujours, que ce qui compte dans un son de porte est le fait qu'il évoque une porte ; en d'autres termes, peu importe le son qu'on choisit tant qu'il s'agit d'une porte. A cela s'ajoutent les précisions du son qui font qu'il est vraisemblable, comme dans le cas de la porte sa matière, la force avec laquelle elle est fermée, etc.

Pourtant, chacun ayant été confronté à la fabrication d'un film a pu remarquer que le choix d'un son lors du montage est souvent motivé en plus par d'autres critères, que

l'on a l'habitude de qualifier de « subjectifs ». C'est d'autant plus remarquable dans le cadre de la musique, où le son d'un instrument peut parfois nous plaire et parfois pas, alors qu'il s'agit du même instrument jouant le même morceau.

De fait, il semble évident que l'apparence du son joue un rôle dans la façon dont le film est reçu par le spectateur, donc que son écoute réduite entre en jeu ; dans quelle mesure alors ?

### **3. Principe et objectif de l'expérience**

Afin de mettre en évidence la façon dont l'écoute réduite influe sur notre perception d'un film, nous avons réalisé l'expérience suivante :

- A partir d'une séquence de *Mulholland Drive* de David Lynch, nous avons d'abord relevé l'ensemble des sons montés en déterminant leurs fonctions respectives ;
- Nous avons ensuite enlevé toute la bande sonore de la séquence, et nous l'avons reconstruite en synthétisant des sons devant jouer le même rôle que ceux mis en évidence dans l'analyse. Nous avons réalisé deux versions : la première est une imitation de la séquence originale, et la seconde est une réinvention ;
- Nous avons ensuite montré ces trois séquences à un ensemble de personnes qui ont répondu à un questionnaire parallèlement au visionnage des extraits, dans plusieurs ordres différents.

La séquence sera décrite en détail plus tard lors de l'analyse, mais nous pouvons préciser dès maintenant qu'elle possède la particularité d'être constituée en majorité de sons non diégétiques, et plus particulièrement des sons abstraits. Ces sons sont liés à des événements visuels, mais ne répondent pas à une logique diégétique. C'est donc un terrain privilégié pour notre expérience, où nous pourrions faire varier ces sons avec une grande liberté.

Une hypothèse que nous proposons est qu'il suffit d'étudier la façon dont sont perçus les sons abstraits, qui ne font pas appel aux écoutes causale et sémantique, pour pouvoir mettre en évidence les tendances générales de l'écoute réduite. Ces tendances seraient

alors, avec cette hypothèse, valables dans une certaine mesure en complément des écoutes causale et sémantique pour l'ensemble des autres cas.

Avec cette expérience, nous cherchons à mettre en évidence le fait que l'esthétique des sons joue un rôle dans la perception d'un film par un spectateur, et cela indépendamment de leurs rôles narratifs. En d'autres termes, nous voulons montrer que ce à quoi ressemble le son a de l'importance.

De manière plus large, il s'agit de montrer qu'il existe une forme de traduction possible entre certaines caractéristiques des sons et les réactions perceptives qu'ils suscitent chez le spectateur, comme les émotions et les sensations.

#### **4. Difficultés et limites**

L'objectif de cette expérience se doit de rester modeste, car le terrain qu'il aborde est particulièrement complexe et délicat, notamment à cause de la pluralité des domaines qui le concernent.

##### *a. L'écoute réduite*

D'abord, l'écoute réduite présente en elle-même une difficulté de manipulation dans le cadre de l'expérience. Schaeffer explique que les caractéristiques des sons mises en évidence par leur écoute réduite ne sont pas des caractéristiques physiques, comme la fréquence ou l'amplitude, et qu'il est essentiel de les en dissocier. Par exemple, il est aisé de remarquer rapidement qu'une voix est rauque, mais il est bien plus compliqué de traduire ce caractère par les caractéristiques fréquentielles du son en lui-même. D'une certaine façon, notre cerveau traduit les caractéristiques physiques de l'onde sonore que nous recevons pour nous permettre de décrire le son avec l'écoute réduite, mais nous ne sommes pas capables de revenir en arrière et d'en reformer ces caractéristiques physiques.

Dans le cadre de l'expérience, cette particularité de l'écoute réduite devient une limite, car nous devons pouvoir mesurer les caractéristiques que nous faisons varier. Par exemple, une voix rauque ne peut être qualifiée expérimentalement de rauque que si ses

caractéristiques physiques sont déterminées, et que nous les associons en général au caractère rauque d'une voix. Et même, de manière très pratique, ce sont des caractéristiques physiques des sons que nous déterminons pour synthétiser un son, et cela est de nature différente avec ce qu'étudie l'écoute réduite.

Ainsi, la première étape est de déterminer ce que nous souhaitons faire varier dans les sons de l'expérience, tout en ayant conscience que cela n'a finalement pas grand-chose à voir avec l'écoute réduite. C'est le mécanisme biologique de l'écoute réduite que nous étudions, mais pas l'écoute réduite en elle-même.

### *b. Limites neuroscientifiques*

Les liens entre l'apparence des sons d'un film et les émotions qu'ils suscitent chez le spectateur gagnent à être envisagés avec le concours d'une approche neuroscientifique.

Les neurosciences cognitives, notamment appliquées à l'audiovisuel, appartiennent à un domaine de la recherche encore jeune<sup>18</sup>. Le principe de cette approche est d'examiner l'activité cérébrale d'un spectateur en train de regarder un film. Cette méthode apporte des données chiffrées, quantifiables, sur la façon dont le film est perçu par le spectateur, permettant ainsi de compléter ce qui est dit par le spectateur lui-même après la projection. L'intérêt réside notamment dans la comparaison des activités cérébrales entre les spectateurs (*inter-subject correlation*), pour mettre en évidence ce qui est partagé par tout le monde.

Par défaut de compétence et parce que cela dépasserait largement le cadre de ce mémoire, nous ne pourrions pas faire intervenir cette approche dans l'expérience.

### *c. Limites anthropologiques*

De la même façon, une approche anthropologique complèterait favorablement cette expérience.

---

<sup>18</sup> “*Neurocinematics: The Neuroscience of Film*”, Uri Hasson, Ohad Landesman, Barbara Knappmeyer, Ignacio Vallines, Nava Rubin, and David J. Heeger, 2008

Les résultats que nous pouvons en tirer ici sont intimement liés à notre expérience personnelle de spectateur, la société dans laquelle nous vivons, l'époque, la culture, etc. Bref, il n'est pas possible de dissocier ce qui est culturel de ce qui est naturel sans la comparaison de résultats entre des individus très différents, autant par leur âge que l'endroit d'où ils viennent.

Le cadre de cette expérience ne concerne donc qu'un groupe réduit d'individus, et gagnerait à être étendu à un groupe plus large et diversifié de personnes.

#### *d. Limites méthodologiques*

Enfin, la dernière limite majeure est méthodologique. Le fait que l'expérience ne soit menée que sur une seule séquence d'un seul film, d'une seule époque donnée, etc., ne permet aucune généralisation des résultats. Tout ce qui est mis en évidence n'est valable que dans ce cas précis, et dans le cadre que nous avons précisé avec les autres limites.

#### *e. Quel intérêt alors ?*

Avoir conscience de l'ensemble de ces limites n'empêche pas l'expérience d'atteindre son objectif : montrer que la perception d'une scène change selon les caractéristiques des sons qui la constituent, et cela indépendamment de leurs rôles narratifs. Ainsi, il faut considérer les résultats de cette expérience comme une justification d'une recherche à venir. Finalement, nous cherchons surtout à montrer ici qu'il existe des liens qu'il faudrait déterminer ultérieurement en adoptant conjointement des approches neuroscientifique, anthropologique, historique et cinématographique.

Plus généralement, le principe de cette expérience pourrait être utilisé dans un contexte autre que celui des sons non diégétiques, pour mesurer par exemple les différences de réception d'un film lorsque ce sont d'autres critères qui sont modifiés, comme les sons diégétiques cette fois par exemple, ou bien la qualité technique de l'image comme du son, l'étalonnage, etc. Il s'agit de voir dans quelle mesure la variation d'un des nombreux aspects du film modifie la perception du film tout entier.

## 5. Caractéristiques des sons

Comme nous l'avons mentionné plus tôt, notre expérience se détache du cadre strict de l'écoute réduite parce que nous avons besoin pour la réaliser de pouvoir quantifier les caractéristiques des sons que nous faisons varier. Plus concrètement, il faut que nous définissions ces caractéristiques de manière suffisamment précise pour pouvoir d'abord les synthétiser, puis ensuite les modifier. Il ne s'agit pas de les mesurer précisément lors de la synthèse, mais simplement de pouvoir les modifier de manière significative afin de voir ce que leur modification entraîne dans la perception du film.

En s'inspirant des critères morphologiques que Schaeffer détermine pour décrire les composantes internes des objets sonores (masse, timbre harmonique, grain, allure, critère dynamique, profil mélodique, profil de masse<sup>19</sup>), et en les confrontant à notre pratique du son en général et plus spécifiquement de la synthèse sonore, nous présentons quelques caractéristiques sur lesquelles nous sommes intervenus pour d'abord imiter et ensuite recréer les bandes sonores de la séquence.

### a. *La masse*

Pierre Schaeffer définit la masse comme « le mode d'occupation du champ des hauteurs par le son »<sup>20</sup>. En s'inspirant de cette définition, nous désignerons par masse ce qui constitue le spectre sonore d'un son, obtenu par transformée de Fourier. Elle correspond donc à la partie continue du son, indépendamment de son enveloppe temporelle et de ses variations dynamiques, et évolue dans le champ des hauteurs selon la fréquence fondamentale du son. La masse d'un son peut être tonique lorsqu'elle fait entendre une hauteur (une note de musique par exemple) ou complexe lorsqu'on ne peut pas désigner de hauteur (un bruissement de feuilles par exemple).

### b. *L'allure*

Toujours inspiré de Pierre Schaeffer, l'allure correspond aux variations dynamiques internes du son, ce qu'on désigne en musique par *vibrato*. Nous élargissons cette notion

---

<sup>19</sup> *Le Son*, Michel Chion, p. 196-202

<sup>20</sup> Cité par Michel Chion dans *Le Son*, Michel Chion, p. 193

en prenant en compte le fait que la variation dynamique du son peut être régulière ou irrégulière, et qu'elle peut couper totalement le son avant de le faire réapparaître. En d'autres termes, un son peut être itératif (comme une série de gouttes de pluie) ou continu (comme un moteur de voiture).

### c. *L'enveloppe temporelle*

L'enveloppe temporelle est une courbe décrivant l'évolution du volume du son au cours du temps. Nous utiliserons l'enveloppe ADSR (*Attack Decay Sustain Release*, en français Attaque Chute Entretien Extinction) qui est la plus commune dans les synthétiseurs. Ces termes correspondent à quatre paramètres qui évoluent dans l'ordre chronologique :

- *Attack* : le temps avec lequel le son atteint son niveau maximal ; c'est avec ce temps d'attaque que le son commence, au moment où la touche du synthétiseur est enfoncée.
- *Decay* : le temps avec lequel le son passe du niveau maximal au niveau qui correspond à celui du *sustain*.
- *Sustain* : le niveau relatif d'entretien par rapport au niveau maximal atteint après l'attaque ; c'est le niveau auquel le son est maintenu tant que la touche du synthétiseur est enfoncée.
- *Release* : le temps avec lequel le son passe du niveau d'entretien à une extinction totale, après que la touche du synthétiseur soit relâchée.

### d. *La hauteur*

Nous désignons par hauteur la zone fréquentielle majoritairement occupée par un son. Typiquement, un son peut être grave, médium ou aigu.

Pour un son de masse tonique, la hauteur correspond à la fréquence fondamentale. Pour un son de masse complexe, en revanche, la hauteur désigne plutôt une zone fréquentielle, plus ou moins large et peut surtout être comparée (par exemple, un moteur de voiture est plus grave que des gouttes de pluie).

#### *e. La réverbération*

Enfin, la dernière caractéristique importante que nous présentons est la réverbération. Nous faisons référence par ce terme à la fois à la taille de l'espace (virtuel) dans lequel le son est émis, mais aussi à la distance relative (virtuelle elle aussi) du micro à la source de ce son.

Un son peut par exemple être émis dans une grande pièce très réverbérante (comme une église), ou une petite pièce peu réverbérante (une cabane en bois). Ensuite, pour faire varier la distance par rapport à ce son, il s'agit de choisir l'équilibre de volumes entre le son direct et le son réverbéré ; en privilégiant le son direct, le micro sera près de la source, et en privilégiant le son réverbéré, le micro s'en éloignera.

#### *f. Le mixage*

Nous ne prenons pas en compte dans cette expérience la question du volume de la source. Nous garderons le même équilibre entre les différents sons que dans la séquence originale. Nous avons vu plus tôt que le mixage est important pour la détermination de la distance à l'histoire du son non diégétique, et il convient donc de laisser cette caractéristique inchangée, dans l'idée toujours de n'étudier que le rôle de l'écoute réduite des sons indépendamment de leurs fonctions narratives.

### **6. Protocole expérimental**

L'expérience est constituée de quatre étapes successives :

- 1) Analyse de la séquence : détermination des sons de la séquence et de leurs fonctions ;
- 2) Montage des deux versions, l'imitation et la réinvention : synthèse des sons occupant les fonctions des sons déterminés dans l'analyse, et variations de leurs paramètres pour la réinvention ;
- 3) Présentations successives des séquences à un ensemble d'individus : les personnes participant à l'expérience remplissent un questionnaire en regardant les séquences ;

#### 4) Analyse des résultats.

Les outils que nous avons utilisés sont :

- le logiciel Pro Tools pour le montage des sons ;
- le synthétiseur analogique Korg Minilogue pour la synthèse sonore ;
- les plug-ins par défaut de Pro Tools en complément de la synthèse précédente ;
- le plug-in FabFilter Pro-R pour la réverbération ;
- le logiciel Izotope RX 6 pour la visualisation du sonagramme des sons originaux et synthétisés.

Le montage son est réalisé en stéréo.

Nous avons pris des libertés esthétiques qui correspondent à l'usage de la pratique cinématographique afin de rendre le plus « humains » possible les différents montages, particulièrement la réinvention, qui est plus libre. Il s'agit de faire en sorte que les montages soient esthétiquement cohérents avec la séquence, afin que la comparaison des visionnages successifs ne soit pas perturbée par l'absurdité de ce que pourrait créer une génération automatique par exemple.

### **7. Analyse de la séquence**

La séquence est constituée de huit plans que nous décrivons ci-dessous, et dure 1 minute et 20 secondes :

- Plan 1 : Deux femmes marchent dans la rue et interpellent un taxi qui s'arrête pour les récupérer. Lorsqu'elles montent dans la voiture, l'image devient floue et la caméra tremble violemment. Fondu enchaîné vers le plan suivant.
- Plan 2 : Le taxi roule rapidement sur un pont, et la caméra le suit avec un rapide panoramique.
- Plan 3 : Les deux femmes sont assises côte à côte à l'arrière de la voiture. Elles ont l'air inquiet.
- Plan 4 : Des images de la ville comme vues à travers la fenêtre de la voiture se succèdent. Fondu enchaîné vers le suivant.

- Plan 5 : Comme le plan 4, d'autres images de la ville s'enchaînent.
- Plan 6 : Même plan que le plan 3, sur les deux femmes dans la voiture.
- Plan 7 : Nouveau plan extérieur, comme les plans 4 et 5.
- Plan 8 : Dans une sorte de parking extérieur, le taxi arrive par la droite du cadre et fait descendre au loin les deux femmes devant une porte. Les deux femmes descendent et se dirigent vers la porte, le taxi redémarre, et au même moment la caméra amorce un rapide travelling avant vers la porte. On voit l'écriteau « Silencio », jusqu'à ce que la caméra atteigne la porte, puis *cut* vers la séquence suivante.



- **Figure 13 *Mulholland Drive***

Nous avons relevé l'ensemble des sons montés dans cette séquence, et cette analyse est présentée dans le schéma ci-dessous :

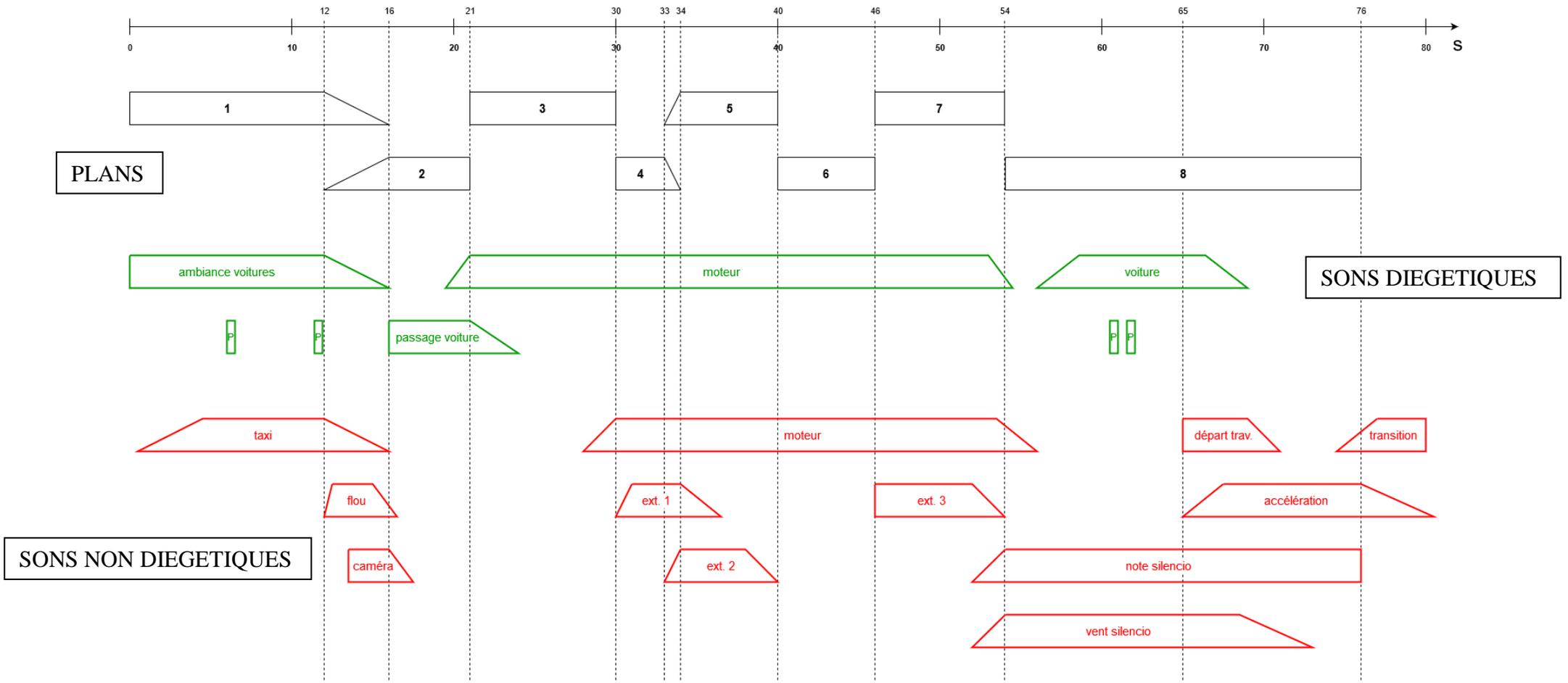


Diagramme analyse de séquence *Mulholland Drive*

En noir sont indiqués les plans, en vert les sons diégétiques, et en rouge les sons non diégétiques. L'axe du temps est marqué par les changements de plans, et également certains points précis au sein des plans qui correspondent à des événements importants dans l'analyse, comme le départ du travelling du dernier plan par exemple.

Nous avons décrits les sons non diégétiques par leur rôle au sein de la séquence, en se détachant au maximum de leur apparence dans la séquence originale. Les seules exceptions sont la « note silencio » et le « vent silencio », qu'il s'agissait de séparer dans l'analyse malgré leurs apparitions conjointes.

Le sonagramme de cette séquence est le suivant (seul le canal gauche est représenté) :

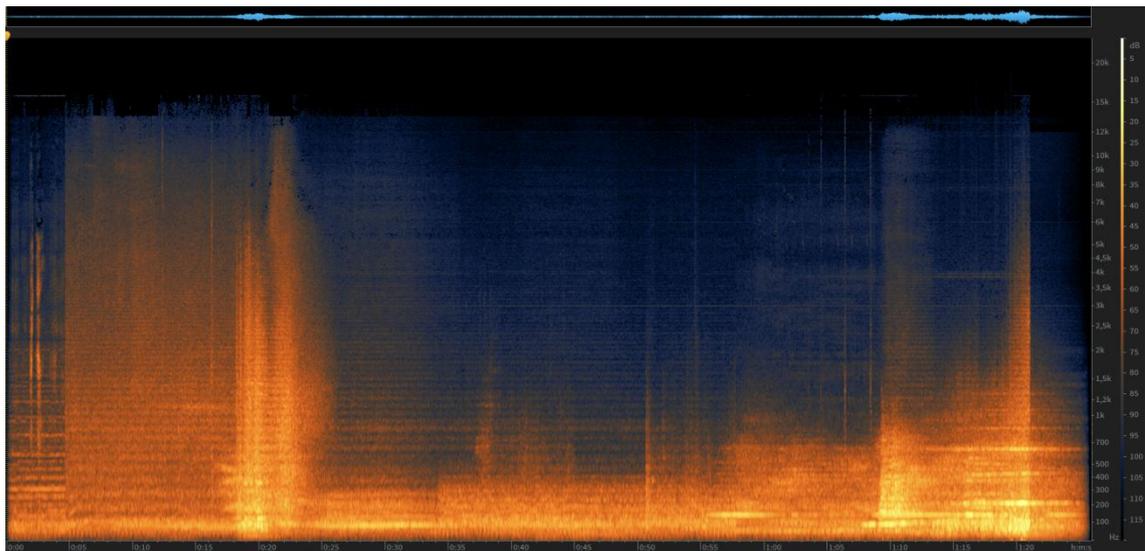


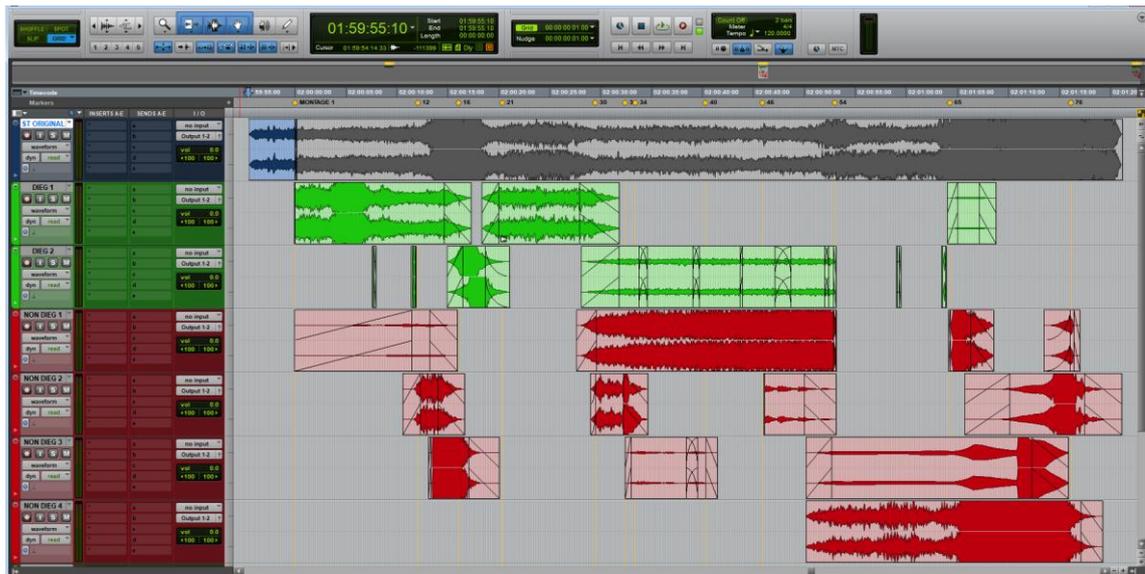
Figure 14 Sonagramme de la séquence originale

C'est à partir de cette analyse et de ce sonagramme que nous avons réalisé les deux nouvelles versions.

## 8. Premier montage : l'imitation

La première version que nous avons recréée est une imitation de la séquence originale. Pour chaque son que nous avons relevé dans l'analyse, diégétiques comme non diégétiques, nous avons tenté en nous fiant à notre écoute et à l'analyse fréquentielle et temporelle du sonagramme de les imiter.

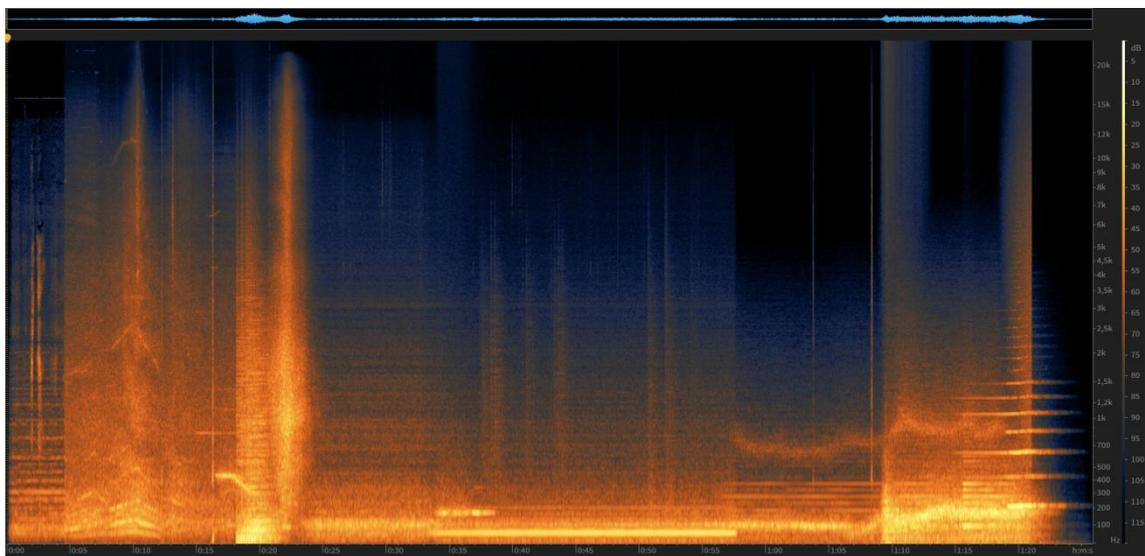
Le résultat est visible dans la session Pro Tools suivante :



**Figure 15 Session Pro Tools de l'imitation**

Comme sur le schéma de l'analyse de la version originale, les sons diégétiques sont en vert et les sons non diégétiques en rouge. La disposition sur les pistes est la même que sur l'analyse originale, et les marqueurs temporels également.

Le sonagramme (canal gauche uniquement) de cette séquence est le suivant :



**Figure 16 Sonagramme de l'imitation**

Il met en évidence les points communs avec la séquence originale, partageant le même squelette et les mêmes évolutions, mais de nombreuses petites différences sont tout de mêmes notables. Cela se comprend aisément, car la reproduction strictement à l'identique nécessiterait une connaissance précise des moyens qui ont été mis en œuvre

pour créer les sons d'origine et un travail de copie minutieux. Mais cela ferait également perdre son intérêt à cette expérience.

L'objectif principal de cette version est d'observer si les petites différences que nous avons mentionnées ci-dessus entre ce montage et la version originale ont une importance ou non dans la perception que le spectateur a de la séquence. En bref : est-ce que cette version et la version originale seront perçues de la même façon ?

## 9. Second montage : la réinvention

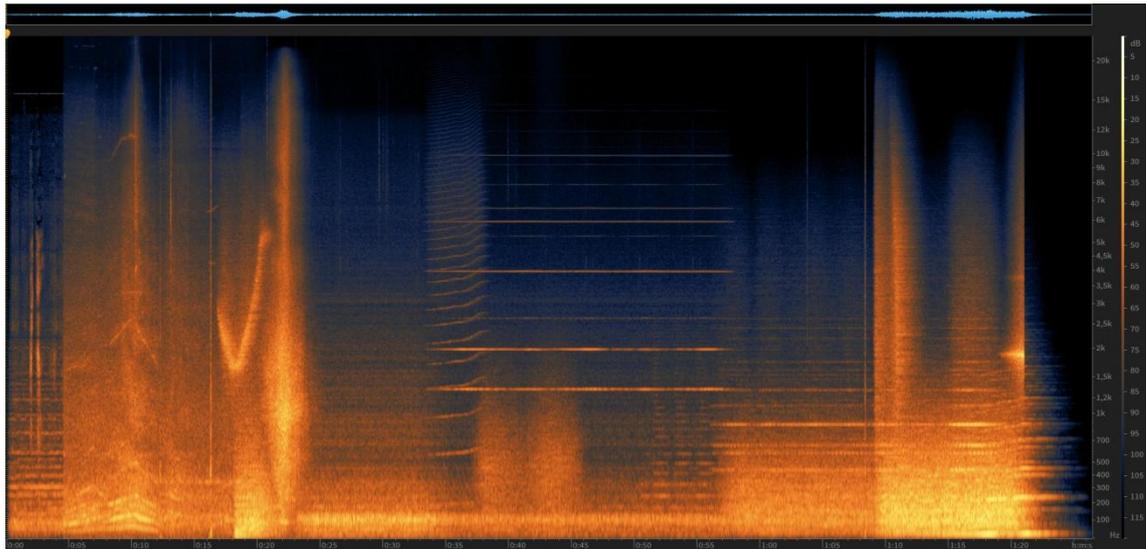
La seconde version est créée à partir de notre premier montage, mais où chaque son non diégétique a été modifié en jouant sur un ou plusieurs des paramètres que nous avons décrits plus tôt (masse, allure, etc.). Les sons diégétiques sont quant à eux exactement les mêmes.

La session Pro Tools correspondante possède donc la même apparence :



Figure 17 Session Pro Tools de la réinvention

Le sonogramme possède par contre des différences notables :



**Figure 18 Sonagramme de la réinvention**

Ici apparaissent de nouvelles composantes, dont les plus visibles sont les tonales. Le squelette de la séquence est toujours le même, cela se remarque avec la correspondance entre la session Pro Tools et l'analyse originale, mais la réalisation pratique est différente.

Les sons non diégétiques sont toujours associés aux mêmes événements, ils occupent les mêmes rôles dans la séquence et sont mixés de la même façon. L'objectif de cette version est d'observer à quel point les différences pratiques de nature des sons montés ont une influence sur la perception de la scène par le spectateur.

## 10. Hypothèses

Cette expérience vise à mesurer l'influence de la réalisation pratique des sons non diégétiques, leur esthétique, sur la perception du film par le spectateur. Les hypothèses qui sous-tendent les deux montages que nous avons réalisés sont les suivantes :

- Le montage original et sa copie se ressembleraient suffisamment pour que leurs effets sur le spectateur soient semblables ;
- Le montage original est techniquement mieux réalisé, plus précis, et devrait se démarquer par rapport à sa copie ;

- Le montage réinventé installe une ambiance différente de l'esthétique du film, particulièrement des images en l'occurrence, et devrait être remarqué comme différent des deux autres.

Nous formulons également les hypothèses suivantes quant aux facteurs qui pourraient influencer la réception de la séquence par le spectateur :

- Sa cinéphilie, c'est-à-dire à quel point le spectateur est habitué à regarder des films ;
- Sa connaissance de *Mulholland Drive*, ou plus généralement de la filmographie, de l'esthétique, de David Lynch ;
- Sa connaissance et sa pratique de la fabrication d'un film, dans le domaine du son particulièrement.

A cela s'ajoute également l'ordre de visionnage des séquences, qui pourrait influencer la réception des séquences par le spectateur. Le premier visionnage, la découverte de la séquence, serait plus pur, tandis que les deux extraits suivants seraient reçus de manière de plus en plus analytique, la surprise et la découverte de l'extrait dépassées.

C'est donc suivant ces hypothèses que nous avons d'abord réalisé six montages différents, que nous avons appelés respectivement versions de A à F, et qui correspondent aux trois séquences (version originale, copie et réinvention) présentées dans des ordres différents. Des liens de visionnage de ces six montages, et l'ordre des séquences de chacun, sont disponibles en annexe.

Nous avons ensuite élaboré un questionnaire que le spectateur remplit au fur et à mesure du visionnage des séquences. Il comporte d'abord des questions préliminaires :

- A quel point vous considérez-vous cinéphile ?
- Avez-vous déjà vu *Mulholland Drive* de David Lynch ?
- Avez-vous une pratique professionnelle de la fabrication d'un film ?

Ensuite sont présents trois champs de description des extraits visionnés, le spectateur mettant la vidéo en pause entre chaque extrait.

Enfin, le questionnaire se conclut par des questions finales :

- A quel point avez-vous trouvé les extraits différents les uns des autres ?

- Avez-vous préféré une version ?
- A votre avis, quel était la séquence originale du film ?

Le modèle vierge de ce questionnaire est disponible en annexe.

## 11. Résultats

L'expérience a été réalisée par 23 personnes aux profils différents. Le détail des résultats est présenté en annexe.

Les principaux critères que nous avons examinés pour juger de l'importance des modifications que nous avons apportées à la séquence sont d'une part la capacité des personnes interrogées à retrouver la séquence originale, et d'autre part la séquence qu'ils ont préférées.

Sur l'ensemble des personnes interrogées, **les trois extraits ont été préférés et considérés comme étant l'extrait original à peu près à parts égales**. L'originale a été un peu plus trouvée que les autres : 35% des cas, contre 26% pour l'imitation et 30% pour la réinvention, 9% ne se sont pas prononcés. L'imitation a été un peu plus préférée que les autres également : 35%, contre 26% pour l'originale, 30% pour la réinvention, 9% ne se sont pas prononcés. 48% des personnes interrogées ont préféré la séquence qu'ils croyaient être l'originale, et 52% ont indiqué deux séquences différentes à ces questions.

L'ordre de visionnage des extraits ne semble avoir aucune influence sur ces résultats. De la même façon, la cinéphilie déclarée des personnes interrogées n'apporte aucun résultat pertinent.

**La connaissance de *Mulholland Drive* et plus généralement de la filmographie de David Lynch joue un rôle important.** Ceux qui n'ont vu aucun film ni extrait de films de Lynch ont pensé en grande majorité (80%) que la réinvention était l'originale, et l'ont préférée en grande majorité également (60%). Les personnes connaissant la filmographie de Lynch ont au contraire tendance à l'écarter : 17% ont pensé que la réinvention était l'originale, et 24% l'ont préférée aux autres.

**La pratique du cinéma semble également jouer un rôle dans la perception des séquences.** Souvent associée à une connaissance de la filmographie de Lynch, les personnes ayant une pratique professionnelle du cinéma ont plutôt cru que l'imitation était la séquence originale. Ils ont également eu tendance à écarter la séquence originale : personne ne l'a considérée comme étant l'originale, et 13% l'ont préférée aux autres versions.

La moitié des personnes interrogées ont trouvé que deux extraits se ressemblaient et que le troisième était différent. L'imitation était considérée parfois comme similaire à l'originale, et parfois comme similaire à la réinvention. L'originale et la réinvention n'ont pas été mises en relation.

Les descriptions des séquences donnent également d'autres indices. La perception de la durée de la séquence est parfois très variable d'un extrait à un autre pour certaines personnes, mais cette durée perçue semble dépendre des préférences personnelles du spectateur. De manière générale, les personnes mentionnant des durées perçues différentes trouvent que la version qu'ils ont préférée était la plus courte.

Le côté musical de certains extraits, la réinvention particulièrement, a été souligné plusieurs fois.

Des ressentis différents de l'ambiance de la scène ont été décrits (mystère, angoisse, magie) mais cela était très variable selon les personnes interrogées. Les séquences préférées sont le plus souvent décrites positivement, tandis que les autres sont décrites comme possédant une ambiance de nature différente (une séquence était mystérieuse et l'autre angoissante par exemple) ou de degré différent (une séquence était moins mystérieuse que l'autre par exemple).

## **12. Interprétations des résultats**

Ces résultats mettent en évidence l'importance de la culture cinématographique du spectateur dans sa perception de la séquence. Connaître l'esthétique de David Lynch était un critère déterminant dans l'appréhension des trois extraits.

Contrairement à nos hypothèses, la séquence originale ne s'est pas démarquée. Pourtant, près de 80% des personnes interrogées ont considéré que les trois extraits n'étaient pas

tous similaires, et ce pourcentage atteint 96% si l'on y inclut ceux qui ont trouvé qu'ils n'étaient que « plutôt similaires ». Cela révèle donc que malgré les squelettes identiques des trois séquences, leurs réalisations pratiques entraînent des perceptions différentes chez les spectateurs.

Le fait que les trois extraits aient été choisis et préférés à parts égales nous permet d'émettre l'hypothèse que le montage original était suffisamment efficace pour qu'un nouveau montage issu de modifications des sons qui le constituent fonctionne malgré tout. Les séquences ne perdent pas en cohérence malgré leurs esthétiques différentes.

Un résultat intéressant à interpréter est la façon dont les personnes ayant une pratique professionnelle du cinéma ont majoritairement choisi l'imitation plutôt que la séquence originale. Nous pouvons émettre l'hypothèse que la qualité technique et l'époque de création de ces séquences respectives entrent en jeu. *Mulholland Drive* date de 2001 et a été réalisé avec les outils de l'époque, dans une esthétique particulière empreinte de la culture du son analogique et de manipulations physiques des sons. Au contraire, l'imitation que nous avons réalisée a été fabriquée avec surtout des outils numériques modernes, avec la même logique de travail que celle à laquelle les personnes interrogées sont confrontées dans leur pratique professionnelle. Cette attirance vers ce que l'on connaît expliquerait pourquoi cette tendance n'est présente que chez les personnes ayant une pratique professionnelle du cinéma et pas les autres. Cette hypothèse est également appuyée par le lien qui a été suggéré par les personnes interrogées entre l'imitation et la réinvention. Leur principal point commun, que ne partage pas la séquence originale, est la personne qui les a fabriquées et les outils avec lesquels elles ont été réalisées.

Ces considérations poussent donc à insister sur l'importance de la culture personnelle du spectateur dans la perception d'une séquence. Et malgré tout, elle n'a pas fait apparaître de version de la séquence comme « meilleure » que les autres. Cela pousse à dire que d'une part l'essence de la séquence n'a pas été bouleversée par les différentes réalisations pratiques de ses sons non diégétiques mais que d'autre part ces différentes versions ont été perçues de manières différentes par les spectateurs, probablement d'autant de manières qu'il y avait de spectateurs.

### 13. Conclusion : écoute réduite et culture

Cette expérience met en évidence deux grandes tendances.

En premier lieu, elle pousse à relativiser l'importance de l'écoute réduite dans la perception d'un film par un spectateur. En effet, l'essence de la séquence que nous avons étudiée n'a pas été altérée par les modifications que nous y avons appliquées, et aucune version ne s'est démarquée par rapport aux deux autres. L'écoute réduite a entraîné des perceptions différentes des séquences, comme si elles avaient toutes des couleurs, des ambiances, différentes, mais le cœur de la narration n'a pas été affecté.

Dans un second temps, l'expérience met en évidence le caractère profondément culturel de la perception du film par un spectateur, et par conséquent la difficulté de maîtrise des effets de l'écoute réduite. En d'autres termes, les émotions ressenties par un spectateur au moment du visionnage d'un film dépendent de son expérience personnelle, et cette corrélation est encore plus forte dans le cas des sons abstraits.

Le lien que nous avons fait entre écoute réduite et sons abstraits est également relativisé par l'expérience. L'attitude du spectateur face à ces sons est si variable et dépendante de son expérience qu'ils peuvent être perçus comme musicaux parfois, ou bien comme des sons réalistes à d'autres moments. Le rapport du spectateur au film dépend alors grandement de son rapport à ces sons abstraits et de sa capacité ou non à identifier leur source.

Comme nous l'avons expliqué plus tôt dans cette partie, cette expérience gagnerait à être approfondie avec le concours d'autres domaines de recherches :

- Les neurosciences : elles permettraient d'examiner les liens entre les caractéristiques physiques des sons du film et les réactions biologiques des spectateurs ;
- L'anthropologie : en complément des neurosciences, elle permettrait de distinguer ce qui dans la perception d'un film relèverait de la culture et ce qui serait inné.

Elle gagnerait également à être élargie à des films plus longs, dans lesquels la narration entre plus explicitement en jeu que dans la séquence que nous avons choisie. Si l'expérience était menée sur un film entier présentant des enjeux narratifs, il serait alors

possible d'examiner la façon dont l'esthétique des sons non diégétiques, toujours indépendamment de la façon dont ils sont montés, modifie ou non la perception de l'histoire par le spectateur.

## Conclusion

Nous avons étudié dans ce mémoire la façon dont les sons non diégétiques contribuent au récit filmique. En tant qu'éléments étrangers à l'histoire et pourtant très largement utilisés dans les films de fiction, les sons non diégétiques interagissent de manière unique avec le reste du film et participent grandement à la façon dont l'histoire est perçue par le spectateur.

Nous avons d'abord établi une typologie des sons non diégétiques en les distinguant par leur rapport à leur source et le type d'écoute spectatorielle qu'ils entraînent. Ainsi, la musique entraîne surtout une écoute sémantique d'un langage musical, la voix parlée surtout une écoute sémantique d'un langage verbal, les sons réalistes surtout une écoute causale, et les sons abstraits surtout une écoute réduite. Les sons non diégétiques appartiennent majoritairement à un type, mais peuvent passer d'un type à un autre au cours d'une séquence, d'un film, voire même appartenir à plusieurs types en même temps.

Le caractère diégétique d'un son n'est pas contenu dans le son en lui-même. La diégèse est une fabrication qui ne préexiste pas au montage. Bien plus vaste qu'une distinction entre les sons directs et les sons ajoutés en post-production, le caractère diégétique des sons suit une logique qui est purement cinématographique. Nous avons ainsi ajouté à la distinction binaire classique tout un ensemble de niveaux intermédiaires que nous avons appelés les sons à caractère non diégétique.

Dans un second temps, nous avons étudié les fonctions narratives des sons non diégétiques. Cela nous a amené à introduire une notion que nous avons appelé la distance à l'histoire. Elle qualifie le rapport d'un son non diégétique à l'histoire. Nous avons dit d'un son qu'il est *proche* de l'histoire lorsque la diégèse semble ne pas pouvoir exister sans ce son. *A contrario*, un son est *loin* de l'histoire lorsque la diégèse semble indépendante de lui, qu'il n'y a aucun doute sur le fait qu'elle existe sans lui.

Cette distance à l'histoire, pondérée par le caractère diégétique des sons, entraîne deux fonctions majeures des sons non diégétiques que sont la narration et la sensation. Les sons proches de l'histoire favorisent surtout la sensation et ont tendance à ne pas être remarqués en tant que sons non diégétiques par le spectateur. Les sons loin de l'histoire favorisent au contraire la narration et s'imposent au spectateur comme étrangers à

l'histoire et éléments du récit, lui faisant par là même prendre conscience qu'on est en train de lui raconter une histoire. Comme pour le caractère diégétique, cette notion dépasse une distinction binaire et se représente par un axe continu allant de sons très éloignés de l'histoire à des sons très proches.

Enfin, nous avons examiné le rôle de l'esthétique des sons non diégétiques dans la perception d'un film, indépendamment de leurs fonctions narratives. Cela s'est traduit par une expérience menée sur une séquence extraite de *Mulholland Drive* de David Lynch, dont nous avons enlevé toute la bande sonore et recréé ensuite deux versions, une imitation et une réinvention.

Cette expérience a mis en évidence l'importance cruciale de l'expérience personnelle du spectateur dans sa perception d'un film, de son caractère profondément culturel. L'effet de l'esthétique des sons non diégétiques est donc intimement lié à des éléments par nature incontrôlables. Mais malgré tout, le cœur de la narration contenu dans le montage est suffisamment important pour que la séquence que nous avons étudiée ne soit pas complètement dénaturée par les modifications de son esthétique que nous y avons apportées.

Cette approche du son non diégétique en tant qu'outil narratif s'inscrit dans une logique de l'analyse cinématographique qui ne sépare pas les images et les sons. Un son ne peut être qualifié de diégétique que parce qu'une diégèse existe, et celle-ci est créée par le montage d'images et de sons. C'est cette association aux images qui fait du son au cinéma un sujet profondément différent d'autres domaines du son. Ce côté multiforme du cinéma est la cause de sa complexité et cette nature ne peut être ignorée, tant dans son analyse théorique que dans la pratique cinématographique.

Ce mémoire m'a été énormément bénéfique car il m'a permis de comprendre tout un ensemble de phénomènes que j'observais devant des films, mais sur lesquels je n'avais aucun appui théorique pour les expliquer. Le film que j'ai réalisé en tant que partie pratique de ce mémoire a soulevé de nombreuses questions auxquelles j'ai pu répondre par cette approche théorique.

La réflexion menée dans ce mémoire gagnerait à être poursuivie en étudiant les raisons qui poussent un réalisateur à utiliser des sons non diégétiques, ou au contraire à ne pas en utiliser. C'est une question de style, de langage. Pourquoi introduire les sons non

diégétiques dans son langage cinématographique ? Une étude historique de l'utilisation des sons non diégétiques, à la lumière des notions théoriques développées dans ce mémoire, pourrait amener des éléments de réponse.

Plus généralement, je ne doute pas que ma pratique du cinéma en sera à présent définitivement changée. De manière très pratique, j'aimerais développer une façon de concevoir le montage son d'un film qui dépasserait la distinction classique entre les dialogues, les ambiances, les effets, les bruitages et les musiques, en se concentrant plus précisément sur la distinction entre les sons diégétiques et les sons non diégétiques. En dehors des considérations techniques qui imposent par exemple de séparer les dialogues du reste de la bande sonore, il me semble que séparer les sons diégétiques des sons non diégétiques nous rapproche du film, du sens profond des séquences, et clarifie tout le processus de fabrication. Cette distinction permet de répondre de manière très pratique à des enjeux importants de narration.

Avant ce mémoire, j'étais profondément convaincu que le son au cinéma dépasse largement le caractère juxtaposé aux images qu'on lui attribue parfois. J'en suis à présent d'autant plus convaincu, et concevoir le montage comme un assemblage indéfectible d'images et de sons me semble être la seule conception valable.

## **Bibliographie**

BESSALET Jean, GARDIES André, *200 mots-clés de la théorie du cinéma*, éditions Cerf, collection Septième Art, 2011.

CHION Michel, *Guide des objets sonores, Pierre Schaeffer et la Recherche Musicale*, éditions Buchet/Chastel, 1994.

CHION Michel, *Le Son*, éditions Nathan, collection Cinéma/Arts Visuels, 2017.

GEAUDREAULT André, JOST François, *Le Récit Cinématographique*, éditions Nathan, collection Fac. cinéma, 1994.

GENETTE Gérard, Frontières du récit. In: Communications, 8, 1966. Recherches sémiologiques : l'analyse structurale du récit. pp. 152-163.  
([https://www.persee.fr/doc/comm\\_0588-8018\\_1966\\_num\\_8\\_1\\_1121](https://www.persee.fr/doc/comm_0588-8018_1966_num_8_1_1121)).

HASSON Uri, LANDESMAN Ohad, KNAPPEMEYER Barbara, VALLINES Ignacio, RUBIN Nava, HEEGER, David J., “*Neurocinematics: The Neuroscience of Film*”, 2008.

JULLIER Laurent, *Le son au cinéma*, éditions Cahiers du Cinéma, collection Les Petits Cahiers, 2006.

## **Filmographie**

ALLEN Woody, *Annie Hall*, 1977

FINCHER David, *The Game*, 1997

GODARD Jean-Luc, *Bande à part*, 1964

HAND David D., *Bambi*, 1942

HSIAO-HSIEN Hou, *Millennium Mambo*, 2001

JARMUSCH Jim, *Dead Man*, 1995

JENNINGS Humphrey, *The Silent Village*, 1943

LUCAS George, *Star Wars*, 1977

LYNCH David, *Mulholland Drive*, 2001

TATI Jacques, *Les Vacances de monsieur Hulot*, 1953

TATI Jacques, *Mon Oncle*, 1958

STRATMAN Deborah, *Village, silenced*, 2012

WEERASETHAKUL Apichatpong, *Oncle Boonmee, celui qui se souvient de ses vies antérieures*, 2010

## **Annexes**

### **(1) Annexe 1 : séquences de l'expérience de la partie 3**

Les liens de visionnage sont valides à la date de rendu de ce mémoire (juillet 2020).

#### **Version A :**

<https://www.dropbox.com/s/8x56awdttlvk8dm/EXPERIENCE%20MULHOLLAND%20DRIVE%20-%20VERSION%20A.mov?dl=0>

Ordre des extraits : imitation – séquence originale – réinvention

#### **Version B :**

<https://www.dropbox.com/s/t4m9kcsbey5w4ow/EXPERIENCE%20MULHOLLAND%20DRIVE%20-%20VERSION%20B.mov?dl=0>

Ordre des extraits : séquence originale – réinvention – imitation

#### **Version C :**

<https://www.dropbox.com/s/ecr6klcwdeugw2o/EXPERIENCE%20MULHOLLAND%20DRIVE%20-%20VERSION%20C.mov?dl=0>

Ordre des extraits : imitation – réinvention – séquence originale

#### **Version D :**

<https://www.dropbox.com/s/x7iszty2g0zgqvf/EXPERIENCE%20MULHOLLAND%20DRIVE%20-%20VERSION%20D.mov?dl=0>

Ordre des extraits : séquence originale – imitation –réinvention

**Version E :**

<https://www.dropbox.com/s/ir9ubh8q8a80hhj/EXPERIENCE%20MULHOLLAND%20DRIVE%20-%20VERSION%20E.mov?dl=0>

Ordre des extraits : réinvention – imitation – séquence originale

**Version F :**

<https://www.dropbox.com/s/ljhbpcwq427jbgc/EXPERIENCE%20MULHOLLAND%20DRIVE%20-%20VERSION%20F.mov?dl=0>

Ordre des extraits : réinvention – séquence originale – imitation



2. Avez-vous déjà vu *Mulholland Drive* de David Lynch ?
- Oui
  - Non, mais j'ai vu d'autres films de David Lynch ou des extraits
  - Non
3. Avez-vous une pratique professionnelle de la fabrication d'un film ?
- Oui, dans le domaine du son
  - Oui, dans un domaine autre que le son
  - Non, uniquement amateur ou occasionnelle
  - Non

**Descriptions des séquences :**

1. Premier extrait :
  
  
  
  
  
  
  
  
  
  
2. Deuxième extrait :
  
  
  
  
  
  
  
  
  
  
3. Troisième extrait :

**Questions supplémentaires :**

1. A quel point avez-vous trouvé les extraits différents les uns des autres ?
  - Les trois étaient très similaires
  - Les trois étaient plutôt similaires
  - Deux se ressemblaient mais le troisième était différent
  - Les trois étaient plutôt différents
  - Les trois étaient très différents

2. Avez-vous préféré une version ?

La première

La deuxième

La troisième

3. A votre avis, quelle était la séquence originale du film ?

La première

La deuxième

La troisième

### (3) Annexe 3 : résultats de l'expérience de la partie 3

#### Statistiques générales

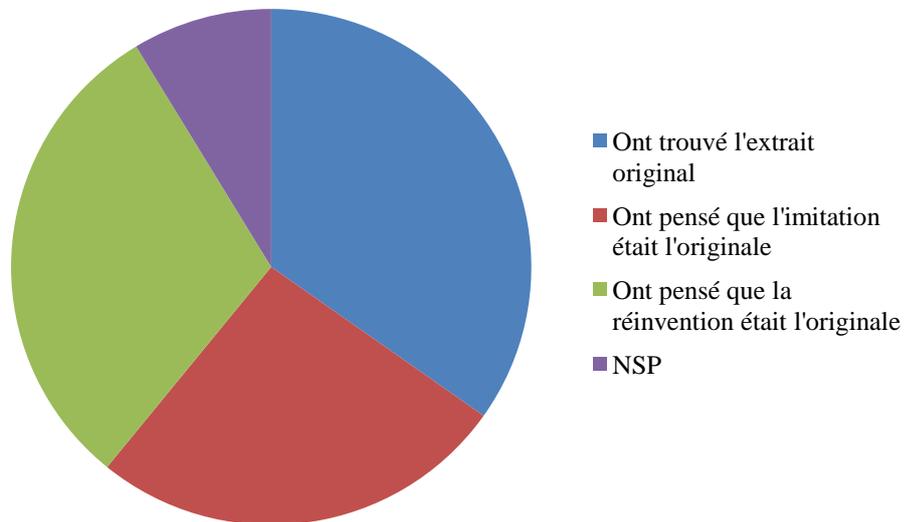


Figure 19 Extrait original sur l'ensemble des personnes interrogées

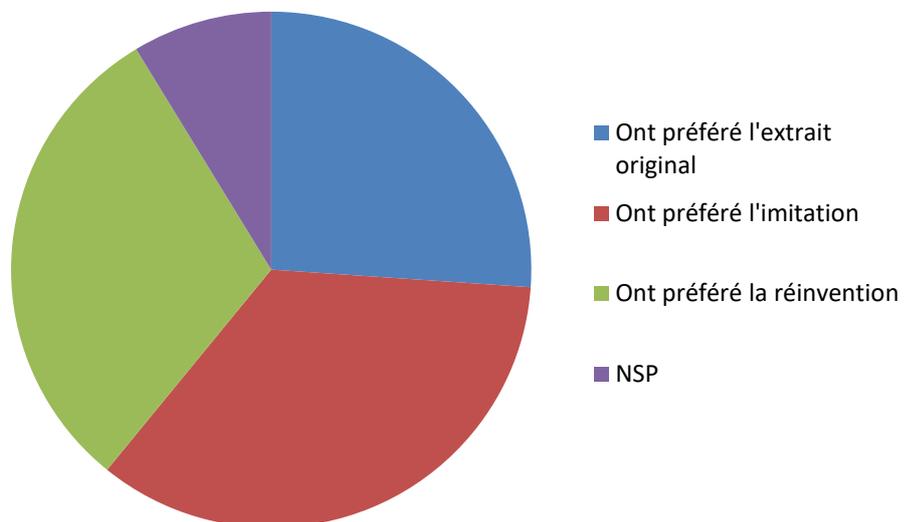


Figure 20 Extrait préféré sur l'ensemble des personnes interrogées

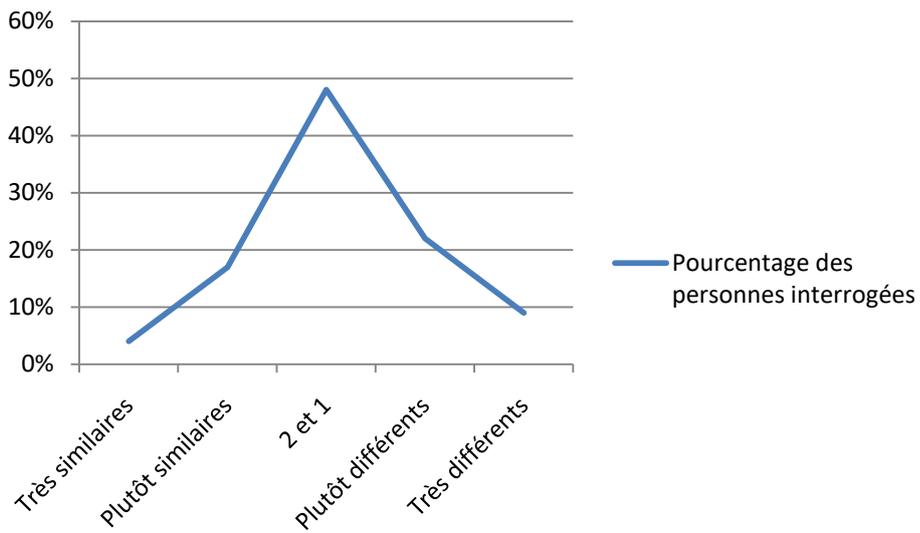


Figure 21 Ressemblances entre les extraits sur l'ensemble des personnes interrogées

### Selon la cinéphilie

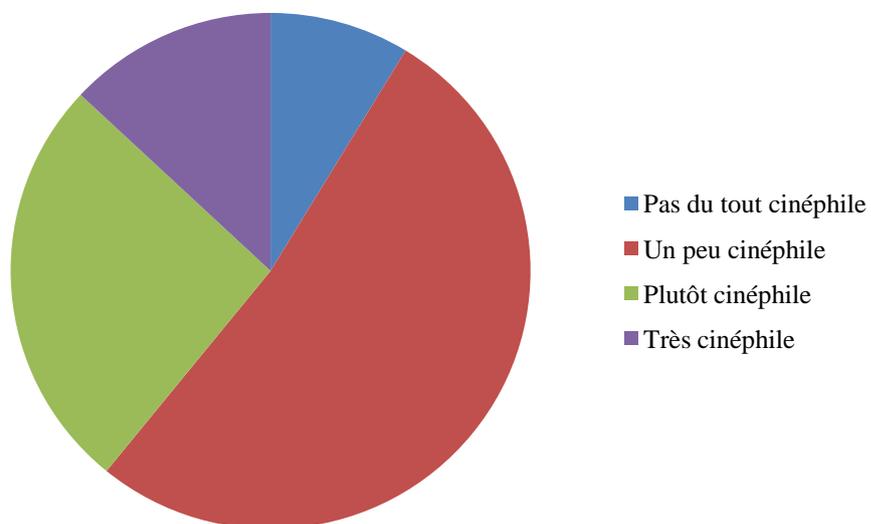


Figure 22 Cinéphilie des personnes interrogées

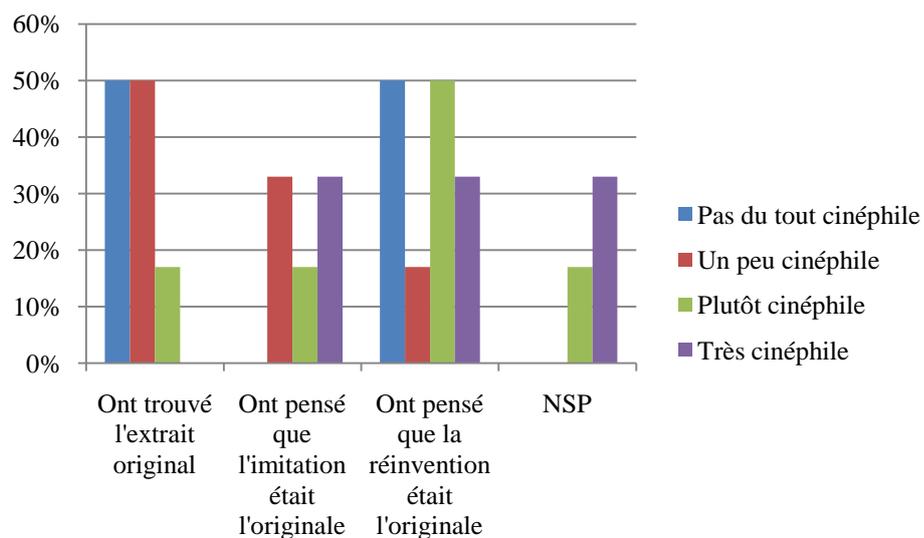


Figure 23 Extrait original selon la cinéphilie

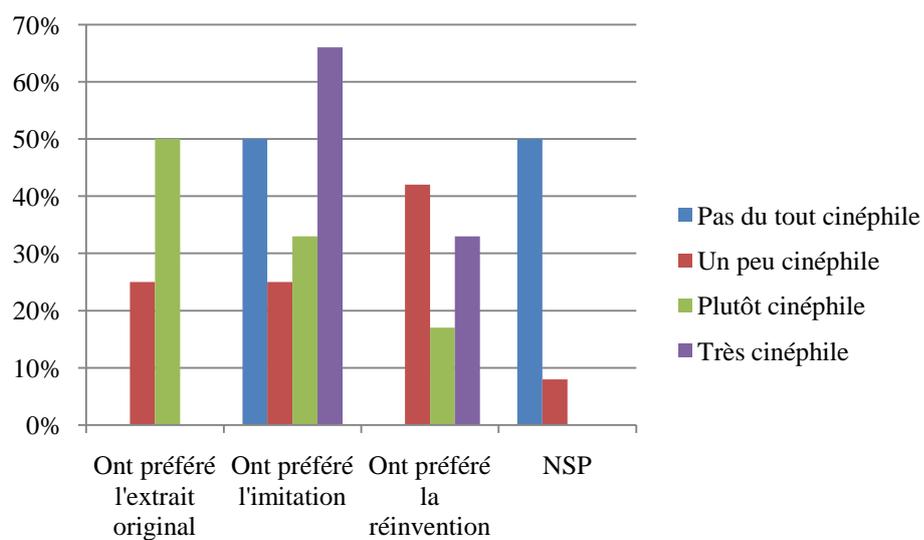


Figure 24 Extrait préféré selon la cinéphilie

Selon la connaissance de *Mulholland Drive* et de la filmographie de David Lynch

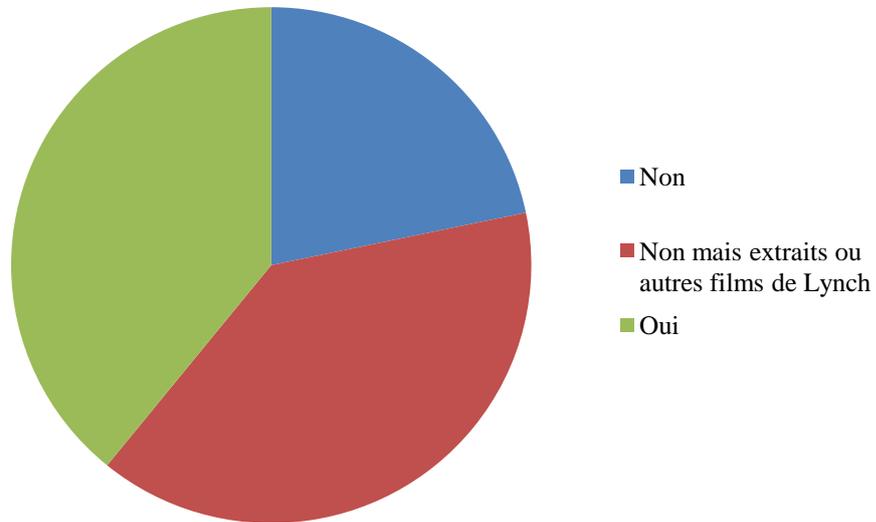


Figure 25 Connaissance de *Mulholland Drive* des personnes interrogées

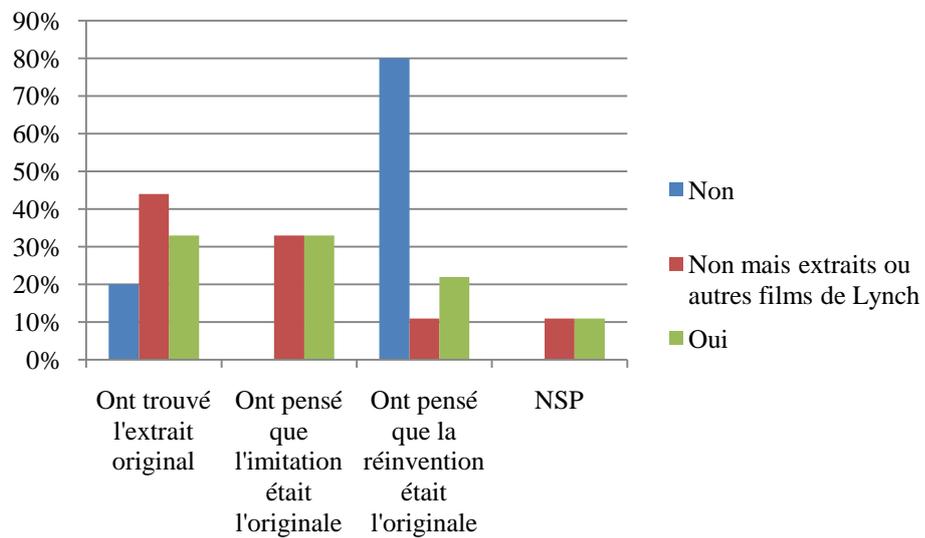


Figure 26 Extrait original selon la connaissance de *Mulholland Drive*

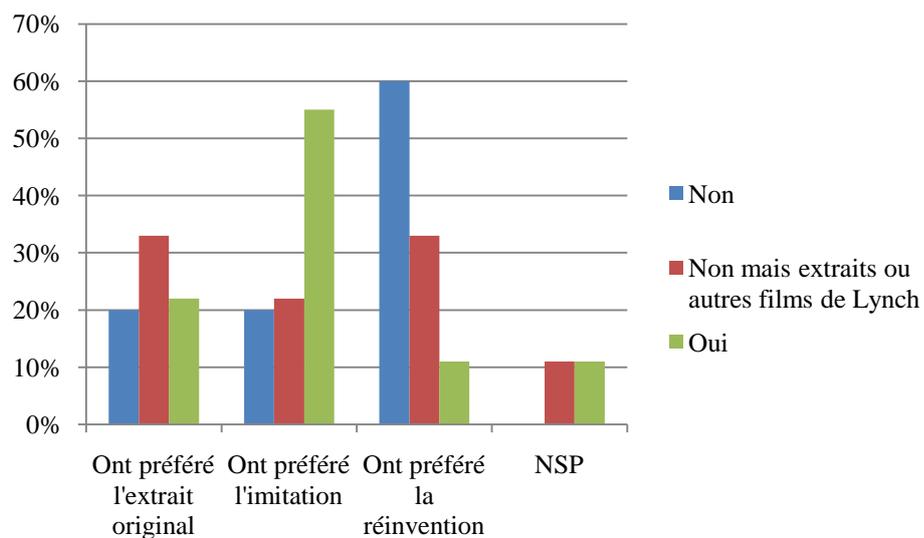


Figure 27 Extrait préféré selon la connaissance de *Mulholland Drive*

### Selon la pratique cinématographique

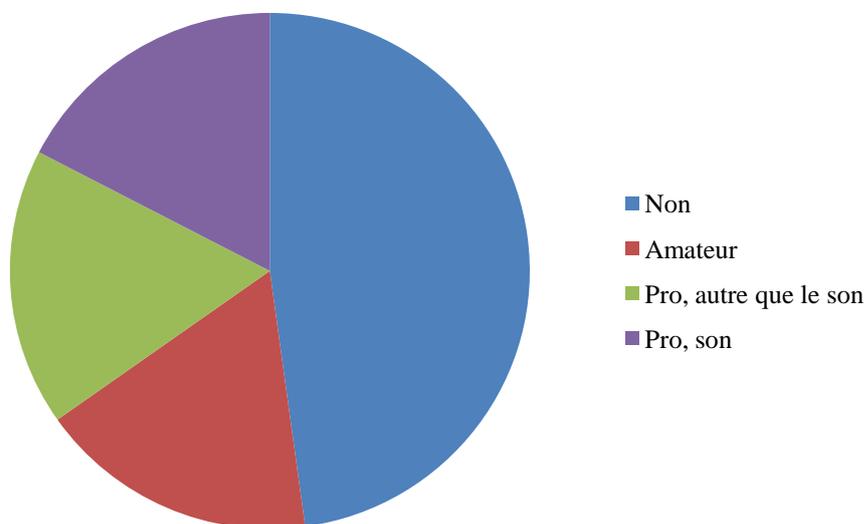


Figure 28 Pratique cinématographique des personnes interrogées

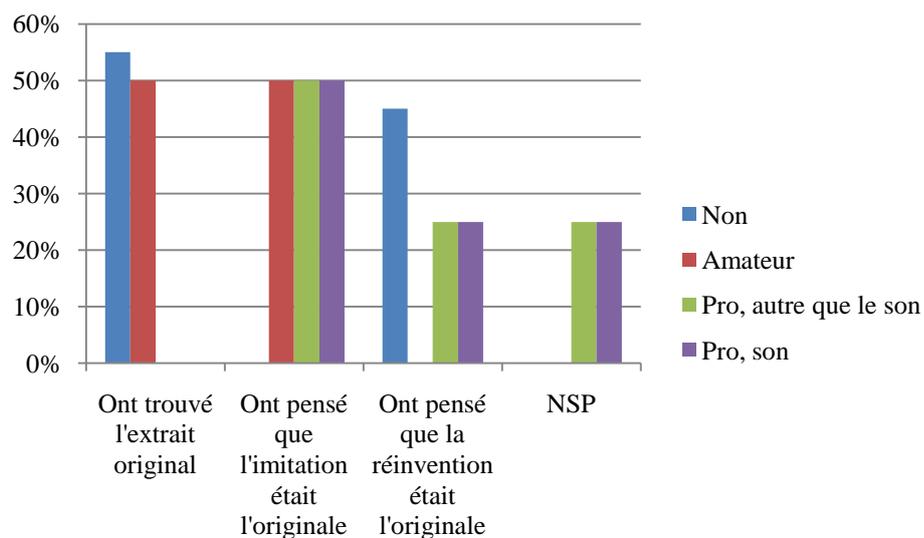


Figure 29 Extrait original selon la pratique cinématographique

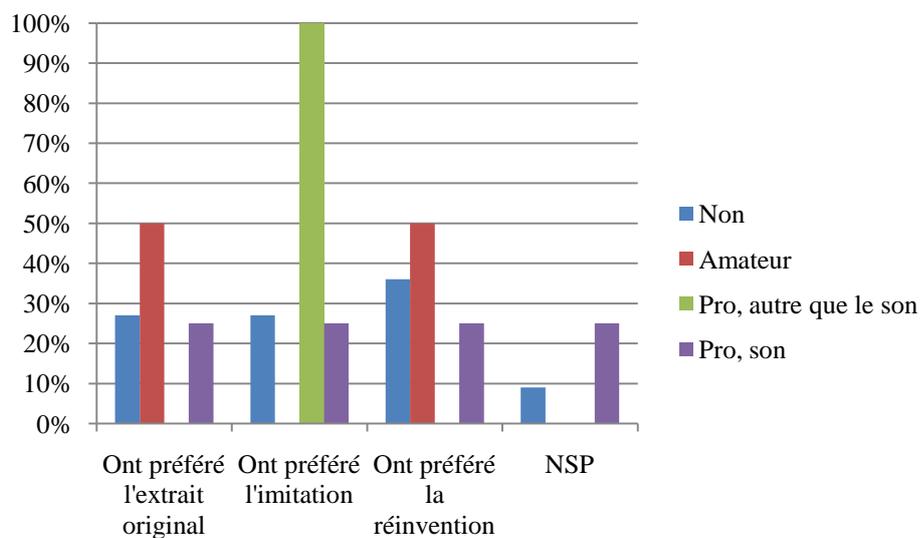


Figure 30 Extrait préféré selon la pratique cinématographique

## Selon l'ordre de visionnage des extraits

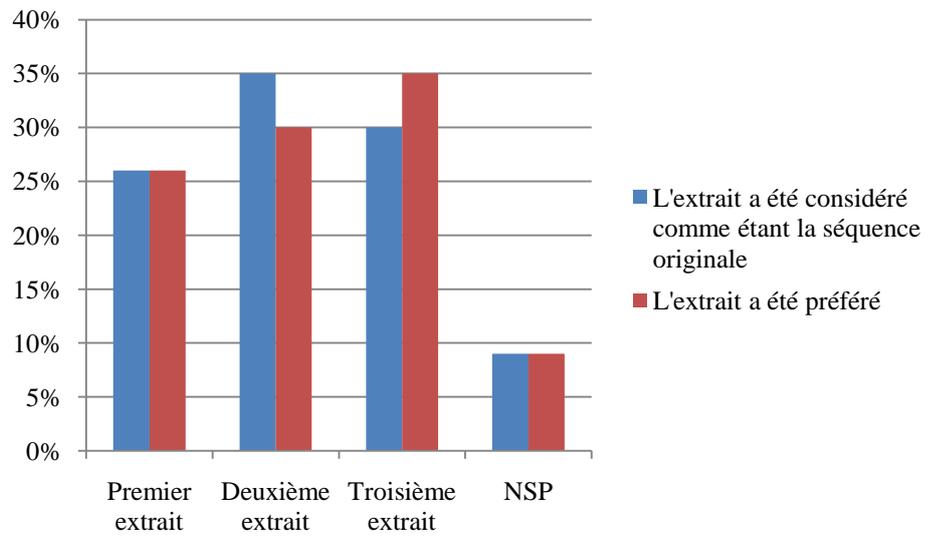


Figure 31 Extrait choisi selon l'ordre

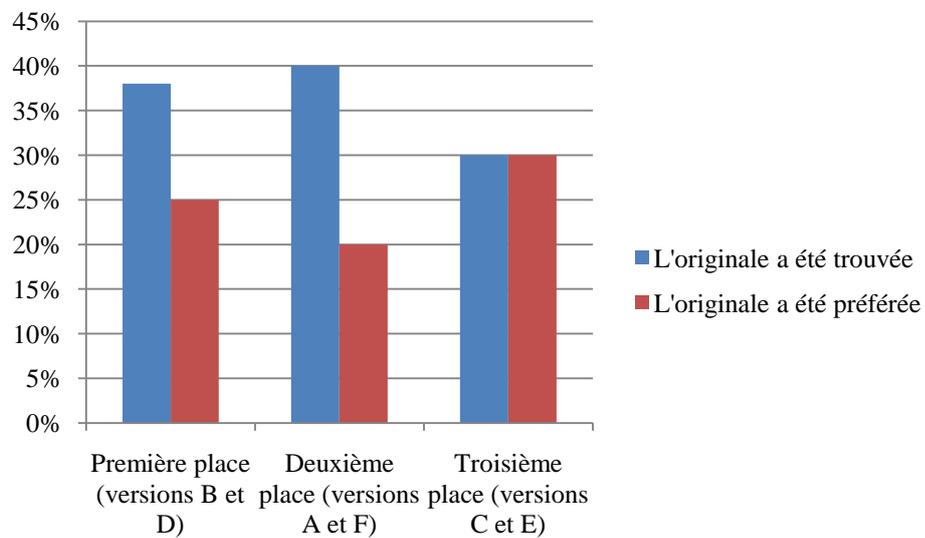


Figure 32 La séquence originale selon l'ordre