

LA PRISE DE SON EN ENVIRONNEMENT BRUYANT

Quelles influences sur le jeu, la mise en scène, la post-production ?

Mémoire de fin d'étude

Yvan Buttaud-Gallot-Lavallée

Département son / Promotion 2025

Sous la direction de Valérie Deloof et Jean-Pierre Laforce

Avec l'assistanat de Jacques Descomps

Sommaire

INTRODUCTION	3
PARTIE I - APPROCHE GÉNÉRALE	5
A - APPROCHE GÉNÉRALE DU TOURNAGE	5
Approches de prises de son	5
Place des sons seuls	8
Ce qui pourrait entraver la possibilité de ce type de prise de son	10
B - APPROCHE GÉNÉRALE DE LA POST PRODUCTION	11
Similitudes dans les façons de travailler	11
Le denoise	13
Contraintes de la version internationale	16
C - ÉTUDE DE CAS	18
PARTIE II - INTÉRÊTS ET INCONVÉNIENTS DE CE TYPE D'ENVIRONNEMENT	24
A - POUR LE JEU	25
B - POUR LA MISE EN SCÈNE	27
C - POUR LE MONTAGE	28
D - POUR LE TRAVAIL DU SON	31
E - POUR LA PRODUCTION	33
PARTIE III - PARTICULARITÉ DES SONS RYTHMIQUES	34
A - CONTRAINTES EN JEUX	34
En décors maîtrisés	34
En décors contraints	38
B - EXPERIMENTATIONS	39
Microphones et directivités	40
Diffusion en basses fréquences uniquement	41
Systèmes de diffusion	43
Denoise de la musique	45
C - SIMILITUDE DES PROBLÉMATIQUES ENTRE LA MUSIQUE ET UN BRUIT RÉPÉTITIF	48
CONCLUSION	50
REMERCIEMENTS et BIBLIOGRAPHIE	52

INTRODUCTION

À l'origine de ce mémoire, le tournage d'une séquence qui rend concret des réflexions alors sous-jacentes. Deux personnages qui parlent au comptoir d'un bar dans une discussion père-fils. On échange en amont avec le réalisateur et nous vient l'envie de faire en sorte que le dialogue soit un peu perturbé par un figurant qui viendrait chercher un verre à côté d'eux, ça ajouterait de la tension dans la séquence. Mais une fois au plateau les choses évoluent, peu de figurants sont dans le cadre et le dialogue est important ; donc on décide, presque comme un automatisme, d'avoir une figuration muette pour préserver la prise de son des voix. On met bien en place le figurant qui vient chercher un verre mais à la première prise le comédien principal ne joue pas du tout avec lui, il n'y a pas de petit regard ou d'interaction, ça devient simplement anecdotique. Dans le même temps, on a du mal à trouver le ton de la séquence, le jeu ne marche pas parfaitement, ce qu'on avait testé au découpage ne fonctionne pas tout à fait, c'est un peu laborieux. Le sujet n'est donc plus d'ajouter notre petite tension à la séquence mais de la faire marcher et ce détail passe naturellement à la trappe. Cependant, à la fin du tournage, je m'interroge. Pourquoi est-ce un réflexe automatique d'avoir un environnement silencieux autour des voix ? À une époque, quand on avait moins de pistes donc moins de micro et notamment pas de micro cravates, moins de dynamique sur bande et peu d'outils de denoise, ça se justifiait probablement davantage. Mais aujourd'hui nos moyens ont évolué et on peut se permettre de faire des prises de son dans des environnements d'une nature davantage bruyante tout en ayant une prise de son des paroles de qualité. D'autant que l'environnement semble influencer le jeu, sa sincérité, l'émotion d'une séquence.

Alors ce choix, qui revient souvent aux chef·fes opérateur·ices du son, de savoir dans quelles conditions de bruit les séquences vont-être tournées, ne doit-il pas être davantage remis en question ? D'autres fonctionnements existent mais semblent plus rares et peu d'expérimentations ont été menées sur ce sujet.

Pour autant, la préservation de la prise de son des voix est centrale, inutile de chercher à ajouter de la sincérité dans le jeu si on ne peut pas le capter. Quelles sont donc les conséquences d'une prise de son en milieu de nature bruyante ? Qu'est-ce que ça influence sur la post-production, sur la mise en scène, sur le jeu ?

Tout au long de ce mémoire, nous nous attarderons sur les enjeux liés aux environnements bruyants servant la narration et la mise en scène et étant donc potentiellement voués à rester jusqu'à la diffusion du film. Il ne sera donc pas question des sons parasites, destinés à être nettoyés. Nous allons par exemple évoquer les enjeux d'un environnement bruyant comme une foule, une fête foraine, un marché, une usine, une boîte de nuit, un port, etc. Mais nous ne nous pencherons pas sur le générateur mis en place par l'équipe, sur le souffle du ventilateur servant à simuler une tempête, ou des avions passants juste au-dessus du plateau dans un film d'époque.

De la même manière, nous nous attacherons principalement à évoquer le travail sur des séquences dialoguées. Et ce, parce qu'enregistrer un milieu bruyant sans la contrainte d'un dialogue se déroulant à l'intérieur pose bien moins de problématiques de prise de son, de postproduction, et encore moins de mise en scène ou de jeu, comme nous le montrerons au cours de ce mémoire.

Ainsi, nous essayerons de nous demander quelles sont les implications d'une prise de son de cinéma dans un environnement de nature bruyante.

Nous chercherons à répondre à cette question en nous attardant sur les enjeux que peuvent induire ce type de prise de son au-delà du travail du son au tournage, mais également pour la postproduction, pour la mise en scène et pour le jeu. Nous ne chercherons pas à donner une manière de faire idéale ou un mode d'emploi, car chaque séquence et chaque film est différent, mais tâcherons de prendre du recul pour faciliter un choix davantage éclairé concernant le type d'environnement sonore au tournage.

PARTIE I - APPROCHE GÉNÉRALE

Attardons nous d'abord sur des approches générales de la prise de son en milieu bruyant. Tâchons de réfléchir aux méthodes de fabrication pour en comprendre les enjeux. Nous aborderons ces points en gardant en tête une des limites de ce mémoire : nous allons parler de généralités sur les films quand toutes ces questions sont avant tout des enjeux de choix au cas par cas, en fonction du film, de la mise en scène, de son esthétique. Tâchons donc de prendre ces différents éléments comme des clefs de lecture qui peuvent nous accompagner dans nos réflexions sur les séquences, qu'il faut voir comme des bases de travail et bien sûr dépasser pour s'adapter à la création ou l'analyse de chaque séquence, de chaque film.

A - APPROCHE GÉNÉRALE DU TOURNAGE

Approches de prises de son

Au cours des différents entretiens et retours d'expérience que j'ai pu récolter, nous pouvons observer des similitudes dans la façon de traiter les environnements bruyant au tournage dans la mise en place technique.

Partons d'une séquence classique où nous avons généralement un dialogue à prendre entouré d'évènements plus ou moins aléatoires qu'il faut pouvoir enregistrer. Nous allons chercher à enregistrer l'environnement bruyant d'un côté, séparé de la prise de son dialogue, et le dialogue de l'autre avec le moins possible d'environnement bruyant à l'intérieur. L'objectif est alors de laisser de la flexibilité en postproduction pour travailler le niveau des voix et le niveau de l'ambiance en fonction des besoins de la séquence. On retrouvera alors classiquement, pour la prise de son des voix d'une part, une ou plusieurs perches et des micros cravates sur les rôles parlants, et pour la prise de l'environnement d'autre part un couple synchrone (LCR ou AB plus généralement).

On essaiera de se rapprocher au maximum de l'action pour gagner en rapport signal sur bruit. Avec la perche d'une part ou en cachant des appoints. Avec les HF d'autre part, qui jouent un rôle essentiel dans ce type de prise de son au vu de leur emplacement constant au plus proche des comédiens. Ils permettent ainsi d'amener du niveau de voix, parfois essentiel pour rapporter une part d'information dans la voix et donc d'intelligibilité dans le cas où de nombreux autres éléments viendraient s'ajouter dans la perche. Ils nécessitent cependant souvent beaucoup de traitement en postproduction, dont du denoise parfois très important, pour avoir ce rôle. Leur avantage d'être dans une

position fixe toujours en proximité implique également de ramener de la présence et de la proximité, ce qui change le rapport au direct – en comparaison à ce que peut offrir la perche – pour un son moins naturaliste. Sans être une solution idéale, il offre donc une sécurité qui paraît nécessaire dans ce type d'environnement. On pourra d'ailleurs chercher à remonter les HF plus proche de la voix quitte à perdre en clarté du timbre mais en gagnant en rapport signal sur bruit. On peut appuyer l'importance des HF en partant du constat que les séquences tournées en milieu bruyant peuvent souvent correspondre à une mise en scène un peu mobile, avec parfois des éléments improvisés. Cet aspect particulier de la mise en scène s'accompagne d'une nécessité de flexibilité dans le dispositif de prise de son pour répondre à ces imprévus. L'utilisation de HF paraît alors d'autant plus importante pour garder un point d'écoute fixe sur les rôles quand le cadre – et donc la place de la perche – peuvent être assez mobiles.

Dans une situation de foule on essaiera classiquement d'avoir la foule comme une masse, d'enregistrer des *walla* (typologie de son dans lequel on entend des personnes parler sans comprendre ce qui est dit), et non comme des individus séparés se démarquant, ce qui viendrait détourner notre attention de la scène qui se déroule et qui est l'élément central. On peut par exemple retrouver ce dispositif dans une séquence de file d'attente devant une boîte de nuit, dans le film *Enzo* de Laurent Cantet et réalisé par Robin Campillo, dans lequel Julien Sicart – chef opérateur son – vient installer un couple à quelques mètres de la file d'attente bondée. La masse de figuration parlant à un niveau assez élevé vient facilement recouvrir les voix des rôles qui sont, pour leur part, prisent avec perche et micros cravates. Cette disposition permet d'avoir une matière stable qu'on peut très simplement utiliser au montage son pour travailler la vie de cette ambiance.

Par ailleurs, ces techniques se retrouvent dans des environnements bruyants maîtrisés comme non maîtrisés. C'est-à-dire dans le cadre de tournage sur lesquels on a la main sur l'environnement sonore, typiquement dans un tournage de fiction avec une équipe de figuration embauchée pour le film ; ou dans des séquences tournées en décors réels, comme un marché en activité par exemple. Dans ces deux cas, on retrouve le fonctionnement perche, HF et ambiance synchrone, adapté selon les situations, complétement d'appoints au besoin.

On peut tout de même retrouver des variations, des techniques qui viennent bien sûr évoluer en fonction des conditions de tournage et des chef·fes opérateur·ices du son. On peut par exemple retrouver l'utilisation de micros à rubans qui réduisent les transitoires,

et peuvent donc être très utiles dans une usine par exemple. On peut aussi évoquer l'utilisation d'un M160 dans *l'Envol*.¹ Sur ce tournage Erwan Kerzanet a utilisé le M160, micro à ruban, couplé avec un autre micro à ruban plus récent, le RE84A. Tout d'abord pour amener un son correspondant au film et à une esthétique franche de l'image, mais aussi pour éloigner la voix du bruit environnant. Comme ce sont des micros avec une sensibilité moindre par rapport aux micros statiques, ils ont le désavantage de demander une proximité importante avec la source et peuvent donc être compliqués à utiliser en fonction des envies de cadre, cependant cette faible sensibilité implique aussi une meilleure séparation des voix de l'environnement sonore.

On retrouve d'autres techniques, parfois plus pratiques. Pour les séquences de manifestations de *Dossier 137* réalisé par Dominik Moll, François Maurel – chef opérateur du son – a fait en sorte que les tops lancés aux acteur·ices ou à la figuration soit fait avec des sons crédibles de manifestations, à l'aide de coups sur des plaques en métal par exemple, plutôt qu'un top vocal qu'il faudrait supprimer ensuite. On retrouve ainsi ce son dans le film final et il permet également d'ajouter une réaction davantage naturelle au jeu.

D'une manière générale, la·e chef·fe opérateur·ice du son pensera sa prise de son en cohérence avec la séquence et le film avec comme attente principale sous-entendu qu'il fasse une prise de son des voix des rôles principaux de qualité. Mais dans le cadre d'une prise de son en milieu bruyant, l'enregistrement de l'ambiance autour de l'action revient aussi de sa responsabilité et doit se faire en cohérence avec la séquence et le film. Ainsi, il pourrait être pertinent de détourner ce fonctionnement classique en fonction des besoins de diffusions par exemple. Un film ayant pour objectif d'être mixé en mono ne devra probablement pas être enregistré de la même manière qu'un film mixé en Dolby Atmos. Les micros pour enregistrer les ambiances peuvent alors varier. On peut par exemple imaginer l'utilisation d'un couple ambisonic mis en hauteur pour avoir une ambiance générale à destination d'un bed en atmos. La place des microphones varie également. Pour un film se déroulant durant une course d'enduro en moto cross, j'ai par exemple équipé d'un HF toutes les motos importantes de manière à ce qu'on puisse travailler leur présence et avoir une continuité dessus, ceci doublé avec une ambiance raccord lors des séquences dialogués pour permettre d'élargir l'image des motos se déplaçant autour des discussions. Cette réflexion doit se faire à partir du

¹ Voir l'article de l'AFSI « Cannes 2022 - Erwan Kerzanet à propos de "L'ENVOL" de Pietro Marcello. »

scénario et, si possible, en lien avec la postproduction, afin de mettre en place une configuration apportant une esthétique pertinente pour le film et une matière suffisante pour le travail en postproduction.

Place des sons seuls

Cette matière ramenée par le tournage passe notamment par l'enregistrement de sons seuls qui restent un élément très important pour construire l'ambiance de la séquence ou du film, particulièrement dans des séquences dans lesquelles on voit à l'image ces éléments bruyants et où on recherchera souvent une densité sonore du lieu. En étant tourné avec un environnement bruyant, nous aurons déjà une base importante de travail, mais il reste essentiel de prendre des sons seuls pour récupérer des éléments spécifiques ou pour permettre une construction en montage son et plus de marge de manœuvre côté spatialisation.

En fonction des séquences, plusieurs types et couches de sons seuls peuvent être nécessaires pour que l'environnement sonore corresponde à la séquence en postproduction. Pour une séquence de foule on pourra par exemple chercher à la fois des walla, mais aussi venir récupérer des précisions avec une ou deux perches assez mobiles. Venir en plus chercher des réactions isolées ou des précisions de groupe en dirigeant la foule peut aussi s'avérer pertinent. En intérieur et en fonction des séquences, il peut aussi être pertinent de faire des sons seuls de déplacement uniquement : faire vivre la figuration mais sans paroles pour habiter la scène. Ces différentes couches vont ainsi pouvoir être superposées pour permettre une maniabilité et une potentielle évolution au sein de la séquence. On pourra ainsi accentuer sur certains éléments, faire taire les voix tout en gardant une présence de la foule avec les sons seuls de déplacements, spatialiser des réactions du public, etc.

Les sons seuls prennent une place d'autant plus essentielle si l'environnement n'est pas constant. Que ce soit parce que la séquence se tourne dans un milieu réel et que l'ambiance environnante évolue au cours de la journée alors que la séquence se déroule dans un temps réduit ; ou qu'il y ait une foule de figurants très bruyants sur le champ alors que seulement cinq personnes sont prévus sur le contrechamp. Je me suis retrouvé dans cette situation sur un tournage cette année. Le court métrage se déroulant dans un bus avec des enfants, nous pouvions créer une ambiance très construite et poussée sur un plan, mais les heures de travail des enfants étant réduite, le plan de travail et le

découpage avaient été pensés pour optimiser les heures de tournage par rapport aux heures de présence et certains plans de cette séquence n'étaient qu'avec quelques personnages alors que le car devait être bondé. Dans ce type de cas, un temps de son seul est essentiel pour raccorder entre les différents plans et faire exister cette ambiance tout le long de la séquence.

Pour préparer correctement ces enregistrements, il semble important que la·e chef·fe opérateur·ice du son se pose la question à la lecture du scénario en réfléchissant aux sons nécessaires par séquence : qu'est-ce qui est le plus signifiant dans la séquence, qu'on ne pourra pas retrouver dans la même qualité plus tard, qui est particulier à ce film ou cette séquence. Cela suppose aussi de commencer, au moins dans une certaine mesure, à réfléchir au son recherché pour le film fini.

L'organisation est souvent la suivante : la·e chef·fe opérateur·rice son essaie de trouver un créneau sur le tournage en collaboration avec la mise en scène, idéalement prévu en amont sur le plan de travail, ce qui permet également à la·e réalisateur·ice d'être présent·e pour l'enregistrement de ces sons seuls, matière souvent très narrative dans des contextes sonores denses. Plus souvent ces sons sont enregistrés sur les pauses ou les samedis et rarement payés, parfois par les assistant·es voir les stagiaires, et très souvent avec un temps très réduit imposant des enregistrements expéditif. On ne peut que regretter ce constat en considérant l'importance de la qualité de cette matière.

Pour autant, on peut aussi retrouver des pratiques permettant de prendre des sons dans de meilleures conditions, d'autant plus quand la majorité du film se retrouve en milieu bruyant. Il est ainsi possible que l'équipe de post-production vienne sur le tournage pour prendre des sons en parallèle du tournage et ainsi d'avoir plus de temps que l'aurait l'équipe de prise de son. Cela permet d'avoir une équipe figuration sous la main ainsi que le décor même s'il peut parfois être compliqué de trouver des créneaux pour faire ces enregistrements au vu des contraintes du tournage. Pour avoir de vraies bonnes conditions de sons seuls, il est aussi possible de revenir après le tournage voir en cours de montage pour prendre des vrais temps de travail permettant de ramener une matière suffisante. Sur des films dans lesquels les séquences bruyantes sont nombreuses, ce type de fonctionnement paraît particulièrement pertinent dans la mesure où l'ambiance sonore vient prendre une place particulièrement importante et demande une matière riche ; il semble alors nécessaire d'organiser de vrais temps de travail consacrés à cette matière.

Ce qui pourrait entraver la possibilité de ce type de prise de son

Tourner avec un environnement sonore dense en direct impose de fait le son de cet environnement dans le rendu final, ou du moins bien davantage qu'un enregistrement en milieu silencieux. Ceci implique donc nécessairement le besoin de cohérence de cet environnement avec la séquence. Que ce soit une cohérence narrative mais aussi une cohérence d'espace, de type de sons, de niveaux, d'imaginaire. Cependant, si des nuisances sonores viennent s'ajouter à cette prise de son, il est évident que l'enregistrement de notre ambiance en direct devient caduc. Le milieu bruyant va certes masquer une partie des nuisances et donc faciliter le travail à ce niveau, mais si ces nuisances surpassent le niveau du bruit que nous enregistrons, il sera d'autant plus difficile de supprimer un avion, des travaux, ou toute autre nuisance sonore ne servant pas la narration.

Puisqu'en tournant dans un milieu sonore riche nous imposons l'utilisation de sons directs de manière plus importante, la qualité de la prise de son est d'autant plus essentielle. Ainsi, les enjeux d'acoustique et de cohérence avec l'espace de la narration ont notamment une place importante. Une acoustique trop grande pourrait rapidement brouiller les sons enregistrés, et créer un brouhaha plus qu'une ambiance définie. De plus, enlever un effet de réverbération sur les voix est certes de plus en plus facile, mais au milieu d'autres sons – voix ou non – lors d'un enregistrement dans un environnement sonore riche, alors ce traitement est bien moins efficace et donne des rendus souvent chaotiques. Le traitement de l'acoustique doit donc être un élément étudié et certains décors pourraient ne pas correspondre en n'étant pas en accord avec les intentions de narration. Dans ce cas, tourner en essayant de rendre silencieux l'environnement facilitera la création d'un espace cohérent en postproduction avec le traitement des directs et l'utilisation de sons enregistrés dans un autre lieu.

Vient avec cette question la notion de repérage. Une présence aux repérages le plus en amont possible paraît ainsi importante pour prévoir les problèmes à résoudre et potentiellement influencer le choix de décor dans un cas trop drastique. Dans des séquences tournées en milieu bruyant, les enjeux sonores prennent une place plus importante que dans d'autres séquences. Il paraît donc plus réaliste de considérer des négociations sur le choix de décor pour que la prise de son soit faite dans de bonnes conditions et que le dispositif offre un rendu de qualité.

B - APPROCHE GÉNÉRALE DE LA POST PRODUCTION

Similitudes dans les façons de travailler

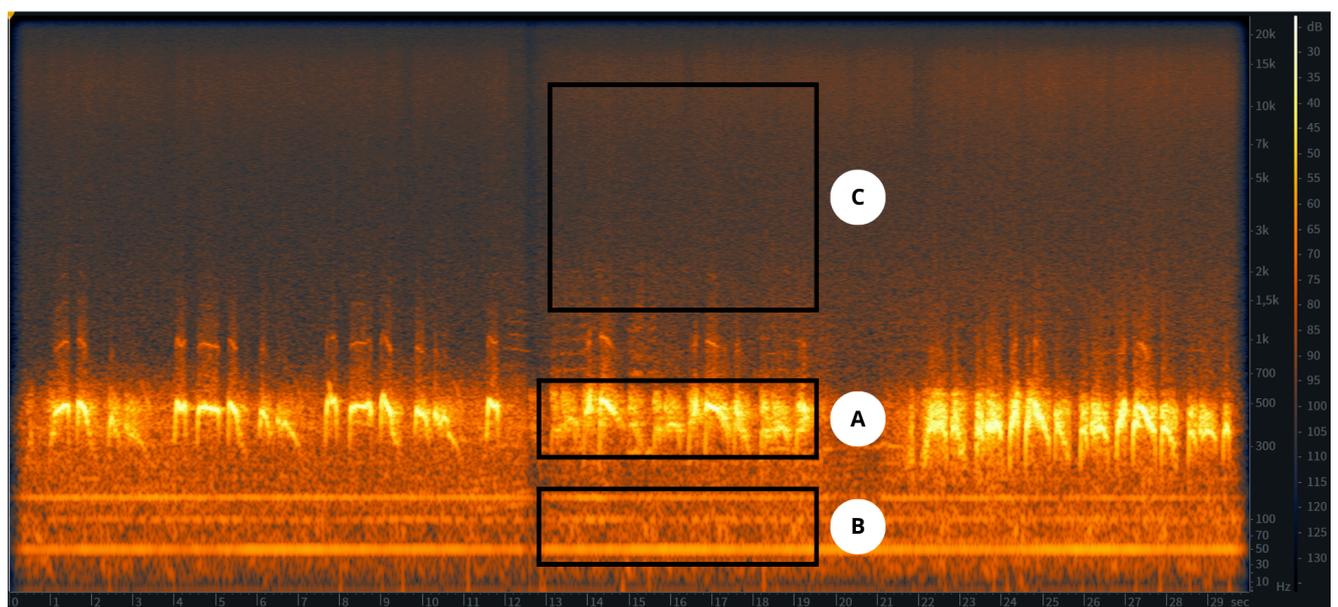
De la même manière que pour le tournage, on retrouve des similitudes dans l'approche générale. Au montage son, on cherchera en premier lieu à travailler les directs en atténuant les éléments qui détourneraient notre attention de l'action principale ou qui limiterait la compréhension. Puis, on pourra ajouter des éléments, en provenance des sons seuls principalement, pour avoir une première base dans laquelle on a peu d'éléments distincts qui ressortent. On viendra ensuite ajouter des précisions avec des sons ponctuels qui donnent de la vie à notre séquence, qui attirent le regard, etc. Ces sons permettent également une spatialisation des éléments une fois arrivés au mixage. Cette approche marche autant pour une séquence de foule, comme celles travaillées sur *Les Misérables* par Matthieu Autin, que des séquences de mer, comme celles travaillées dans *Portrait de la jeune fille en feu* par Valérie De Loof. L'importance du travail des ponctuels pour rajouter de la précision est appuyé par ces deux monteur·euses son comme axe important de la narration.

Une des particularités des ambiances sonores très fortes c'est aussi l'importance que peuvent avoir les ambiances synchrones raccord et la matière qu'elles offrent (ambiances enregistrées en même temps que le direct – en monophonique ou en multicanal – par d'autres micros que ceux captant les dialogues principaux). Dans une séquence plus calme, les ambiances raccords peuvent être utiles pour suivre les évolutions du lieu autour du tournage, pour avoir quelque chose qui corresponde particulièrement au décor puisqu'elles sont enregistrés, de fait, directement sur le décor de manière synchrone. Dans des configurations plus habituelles, les lieux de tournages n'étant pas toujours – voir rarement – en correspondance exacte avec l'ambiance recherchée par la séquence, ces ambiances étant d'une durée réduite (puisque l'équipe travaille entre les plans, l'ambiance n'est souvent exploitable que sur la durée du plan), et les dialogues des personnages étant parfois audibles dans ces ambiances raccords ; ces ambiances se retrouve dans les faits rarement utilisés – bien que les pratiques varient. Pour autant, lorsqu'on tourne dans un environnement sonore plus dense, ces ambiances raccords deviennent une matière incontournable. D'abord l'ambiance du décor se lie avec les intentions de la séquence et les évolutions possibles, mais en plus,

l'environnement est tellement dense qu'on peut facilement exclure les voix des comédiens et les bruits de l'équipe ; ce qui donne une matière exploitable assez riche. De plus, avoir ces ambiances synchrones permet aussi de renforcer l'environnement sonore qui rentre dans la perche, notamment en permettant une diffusion multicanale. Par ailleurs, à chaque changement de plan, on prend le risque qu'un son fort et notable commence juste avant la coupe et continue après celle-ci. Au montage son on va donc chercher à faire continuer ce son sur le plan suivant de manière à ce qu'on ne ressente pas la coupe d'ambiance et qu'on ait l'impression d'une continuité. Si cette coupe se passe au milieu d'un dialogue, il peut très vite être difficile de faire continuer ce son avec uniquement la perche puisqu'il y a de grandes chances qu'une réplique arrive. Ainsi, l'ambiance synchrone permet aussi de travailler facilement tous ces raccords. Ceci demande cependant une relative proximité avec l'action : mettre l'ambiance raccord dans un coin, sans qu'elle ne bouge, ne permettrait pas ces raccords en étant trop loin de l'action. Cependant, ces logiques de fonctionnement sont à adapter à chaque séquence. Ça peut fonctionner très bien dans un environnement particulièrement bruyant comme une fête foraine ou un marché. Mais dans d'autres contextes moins denses, une ambiance raccord fixe pourrait être plus pertinente. C'est le cas des séquences d'amphithéâtre dans *120 battements par minute* de Robin Campillo, dans lesquelles les personnes présentes dans l'amphithéâtre réagissent beaucoup et il y a beaucoup de sons de foule sans que ça soit un brouhaha constant. Julien Sicart, chef opérateur du son, a placé un couple LCR au plafond de cet amphithéâtre ce qui permet de renforcer l'ambiance de la pièce sans être trop collé à l'action, de faire sentir la foule dans cet amphi, ce qui correspond avec la séquence et les envies de mise en scène. Dans tous les cas, l'ambiance raccord paraît donc essentielle dans un environnement sonore bruyant mais elle peut être de différentes formes en fonction du film, dans son placement comme dans son nombre de canaux.

Le denoise

Tourner en milieu bruyant c'est aussi accepter d'être moins en contrôle : on est en contrôle quand on tourne dans un environnement silencieux nous permettant de faire une belle prise de voix. On contrôle les différents éléments, il est plus facile d'être certain-es de l'intelligibilité de nos personnages, de la marge de manœuvre en postproduction, de la place qu'à le fond, du rapport signal sur bruit. Nécessairement, quand on ajoute une figuration parlante – allant de quelques personnages à une foule –, on est moins en contrôle d'une phrase qui pourrait ressortir. Quand on tourne dans une pièce avec des machines en marche, on peut moins facilement savoir s'il n'y a pas une machine qui va faire un bruit un peu trop fort au mauvais moment. On est moins en contrôle, on prend plus de risques. Mais ces risques sont aussi aujourd'hui beaucoup plus aisément excusés par le développement des outils de denoise (c'est-à-dire des outils permettant un nettoyage des prises de son). Là où il y a quelques années, on travaillait surtout avec de l'égalisation ou de la compression par bande, qui limite les possibilités d'actions, aujourd'hui on a largement plus de marge de manœuvre. Cette évolution est d'abord arrivée avec l'apparition du logiciel RX de Izotope, apparu en 2007 mais plus largement utilisé à partir des années 2013 (avec la sortie de RX 2). Souvent décrit comme le Photoshop du son et permettant une visualisation des sons selon trois axes : en abscisses le temps, en ordonnée la fréquence et en luminance l'intensité (avec, par défaut, du jaune pour les forts niveaux jusqu'au bleu/noir pour les faibles niveaux).



Son de pigeon ramier visualisé sur RX10 de Izotope.

On peut facilement voir sur ce son que le chant du pigeon ressort dans le cadre A (avec quelques harmoniques au-dessus), que des fréquences basses constantes sont dans le cadre B et qu'un bruit de fond est dans le cadre C. En visualisant ces éléments on peut donc facilement, et grâce à tous les outils inclus dans le logiciel, traiter les éléments qu'on considère gênant pour les réduire ou les enlever. Évidemment la difficulté dépend des sons et ça n'est pas non plus un outil magique qui permet de tout traiter.

Avec RX Izotope sont apparus d'autres outils, fonctionnant aussi avec des algorithmes comme Cedar audio ou Acon digital par exemple.

Depuis 2023 on voit apparaître de nouveaux outils de denoise, fonctionnant cette fois-ci avec une base d'intelligence artificielle. Ces outils permettent du nettoyage beaucoup plus simplement – mais aussi avec beaucoup moins de contrôle.

En fonction des outils, on observe différentes spécificités avec des plugins. On peut par exemple citer *Hush* qui tend à être meilleur que les autres disponibles dans la diminution de la réverbération d'une pièce quand *Clear* n'est pas très bon dans ce domaine mais meilleur sur le traitement des bruits d'objets. On a également des différences de fonctionnement, *Clear* ne fonctionne qu'avec trois boutons : "ambience", "voice" et "voice reverb". Quand d'autres outils comme *Clarity* offrent aussi une possibilité de niveau de traitement par bande de fréquence.



Interface du plugin Clear de Supertone

Pour certains éléments sonores – mais pas pour tous – ces outils sont beaucoup plus efficaces que les outils par algorithme et permettent donc davantage de marge de manœuvre. Dans l'utilisation un mélange des outils par IA et par algorithme apparaît encore comme essentiel.

Pour souligner l'importance de ces outils dans le travail aujourd'hui, Rym Debbarih-Mounir, monteuse son, nous explique « On a vu une révolution à l'arrivée de RX Izotope dans notre métier. Avec l'arrivée de l'IA c'est plus qu'une révolution c'est une autre manière de penser. ». Dans des cas avec une configuration plus légère, sans HF, elle utilise notamment ses outils pour créer des “faux HF” en nettoyant beaucoup la perche et en l'ajoutant par dessus le direct pour permettre un regain en intelligibilité quand c'est nécessaire sans dénaturer la perche pour autant.

Cependant, il nous manque encore une étape dans ces outils par IA. Aujourd'hui, on a un traitement de “l'ambiance” dans sa globalité, soit tout ce qui est autre que la voix et donc sont inclus les ambiances générales ainsi que les manipulations d'objets, les actions des personnages. Il nous manque encore cette différenciation entre les éléments autres que la voix qui peut être très utile. Dans certains outils on peut également retrouver la séparation voix-musique. Mais il paraît évident que ces outils vont continuer à se développer en ajoutant notamment ces différenciations et deviendront de plus en plus essentiels. Dans le même temps, en permettant davantage de manœuvre, ils permettront aussi davantage de possibilités et de risques au tournage.

On remarque ainsi que ces outils de denoise prennent une place essentielle dans la facilité de traitement des prises de son en milieu bruyant. Cet élément fait donc partie des facteurs qui facilitent ce type de prise de son par rapport aux possibilités que nous avons il y a quelques années. Et cette facilité de traitement se retrouve dans la majorité des environnements bruyants. Quand il y a des fonds très présents, des machines qui font du bruit, etc. les outils par IA sont très efficaces. Quand il s'agit de foules, comme il s'agit également de voix humaines, il est plus difficile de dissocier les voix des personnages principaux des voix de la figuration. Pour autant, on peut tout de même s'en sortir avec des outils comme RX ou en allant chercher dans des doubles si des voix de la figuration viendraient ressortir un petit peu trop. C'est notamment la façon de travailler de Valérie Deloof sur la séquence de *Enzo* – de Laurent Cantet et réalisé par Robin Campillo – dans une séquence de dialogue avec des personnages passant dans une file d'attente de boîte de nuit dont nous avons parlé précédemment, où le retour de

Matthieu Autin sur les films *Athena* ou *Les Misérables* qui ont beaucoup de séquences de foules.

Pour autant, ces outils sont aussi à double tranchant et il faut faire attention à leur utilisation. Tout d'abord, ils peuvent réduire aussi des fréquences de la voix des personnages principaux et réduire notamment les fondamentales ce qui altère le timbre². Il y a donc souvent un choix à faire entre attaquer la voix et réduire le bruit gênant. Par ailleurs, on peut vite avoir tendance à vouloir supprimer tous les petits bruits comme les bruits de bouches par exemple.

« Le plus gros travail avec ces outils c'est de ne pas enlever l'émotion au passage. Et l'émotion ça passe aussi par plein de petits accidents. Il faut faire attention à ne pas trop nettoyer. »³

Contraintes de la version internationale

La construction de la Version Internationale (VI), c'est-à-dire une version du film sans paroles distinctes à destination du doublage dans d'autres langues, peut être vue tout d'abord comme une contrainte importante à ce type de configuration. Particulièrement avec les normes des films et séries dites de plateformes (produit par les plateformes principales comme Netflix, Prime Video, Apple TV, Disney plus) ont des normes très contraignantes sur ce sujet et demandent de retrouver exactement les mêmes sons dans la VI et dans la VO⁴. Ainsi, si une porte claque dans le direct sur les dialogues dans les directs, il faudra la supprimer de la piste dialogue pour, soit en ajouter une nouvelle à mettre dans les effets, soit extraire cette porte du direct pour la mettre dans une piste VI. Mais il est exigé d'avoir strictement le même son dans les deux versions. Pour les films de plateforme cela peut donc être une vraie contrainte en fonction du type de milieu bruyant.

² Thèse de Jérémy Marozeau, *L'effet de la fréquence fondamentale sur le timbre*. ;

³ Entretien avec Rym Debarh Mounir

⁴ **Netflix** : « Ambiences and foley fill should be included to correctly match the original mix. Any non-dialog sound present in the main mix must also be represented in the M&E mix. »

Amazon Prime vidéo : « All M&E tracks should be fully-filled and include all music (Source, Score, etc.), sound effects (SFX), room tone (Ambience/Backgrounds) and Foley as they occur in the domestic mix. Any production effects (PFX) that overlap with dialogue must be replaced/covered with cut SFX or Foley.

Note: Non-fully-filled M&E files will automatically be rejected for fixes. »

Sur une séquence de foule par exemple, ça ne fait pas beaucoup de différences. Même si l'ensemble des directs risque d'être inutilisable car contenant des dialogues en français, de toute manière il faudra refaire toutes les voix de la figuration dans les autres langues – mise à part quelques walla qui peuvent évent être ajoutés dans tous les cas en option – et les quelques bruits du directs qui manqueraient seront facilement refaits en bruitage. Si la séquence avait été prise dans un environnement silencieux ça aurait permis d'éviter quelques bruitages mais qui paraissent – de manière générale – demander une quantité de travail assez minime. D'autant que des sons seuls de walla sont généralement ajoutés pour venir supporter le direct et peuvent donc rester simplement en option pour la VI. Pour autant, dans des séquences où l'environnement sonore est d'une autre nature que de la voix, cet enjeu peut devenir plus gênant. Dans une séquence tournée dans un garage automobile par exemple, on peut imaginer beaucoup de sons mécaniques ou de moteurs qui viendraient se superposer aux voix. Dans ce cas, le fait d'extraire ces sons du direct – ce qui est de plus en plus facile grâce aux outils de denoise par IA – paraît le plus cohérent mais demande un travail supplémentaire. Pour les sons au premier plan ce travail semble, dans une globalité des cas, assez aisément réalisable. Le problème se pose davantage pour tous les sons d'arrière-plan, moins présents dans le direct et se confondant avec la voix. À l'heure actuelle, et même si nos outils se perfectionnent rapidement, il est compliqué voire parfois impossible de les extraire. Il faudrait donc les effacer du direct, ce qui est plus facilement réalisable, pour en remettre de nouveaux. Ce principe renforce l'importance de prendre de la matière au tournage, pas seulement pour renforcer les off mais aussi pour combler les trous créés par le travail sur la VI. Cette matière peut être ramenée soit par un travail de sons seuls – dont la quantité doit être prévue en fonction des besoins de la VI –, soit par l'ajout de micros permettant d'exclure les voix et de ne garder que les off, si le lieu le permet. Il est ainsi important d'y réfléchir suffisamment en amont car si cette approche pourrait ne pas être essentielle dans le travail de la VO, elle devient inévitable pour le travail de la VI. D'autant, donc, que cette VI va directement contraindre la VO dans le cas des films de plateforme.

Nuançons pour autant ce propos en rappelant que les habitudes de travail en ce qui concerne les films à destination du cinéma sont bien moins strictes. Des sons de la VO peuvent donc être similaires dans l'intention mais pas nécessairement exactement identiques à ceux de la VI. Il est ainsi possible de monter des sons pour la VI qui ne

seront pas joués pour la VO et inversement. Il est pour autant nécessaire d'avoir, même dans ce cas, une matière suffisante pour mettre en place ce travail.

C - ÉTUDE DE CAS

Maintenant que nous avons commencé à voir les différents enjeux que représente la prise de son dans un environnement bruyant d'une manière générale, attardons-nous sur un cas concret. Nous tâcherons ainsi de nous intéresser aux méthodes mises en place pour s'adapter à une mise en scène et aux contraintes d'un film afin d'entrevoir des idées de fonctionnement.

Nous allons ainsi nous pencher sur le film *Athena*, réalisé par Romain Gavras et paru sur Netflix en 2022, dont voici le pitch du film :

*Rappelé du front à la suite de la mort de son plus jeune frère, décédé des suites d'une prétendue intervention de police, Abdel retrouve sa famille déchirée. Entre le désir de vengeance de son petit frère Karim et le business en péril de son grand frère dealer Moktar, il essaye de calmer les tensions. Minute après minute, la cité Athena se transforme en château fort, théâtre d'une tragédie familiale et collective à venir. Au moment où chacun pense avoir trouvé la vérité, la cité est sur le point de basculer dans le chaos...*⁵

On peut considérer que ce film est particulièrement bruyant. On est souvent dans des séquences de foule avec une équipe de figuration particulièrement nombreuse (avec des séquences à 500 figurant-es), des tirs de mortier et de feux d'artifices, des moteurs, beaucoup de cris, beaucoup de mouvements. Il constitue donc un cas très intéressant à étudier pour voir quelles méthodes ont été mises en place par l'équipe son pour relever ce défi. Cela nous permettra ainsi d'avoir un exemple concret d'une des méthodes de travail qui peut être mise en place dans ce type de séquence.

Ce film a également la particularité de contenir de nombreux plans séquence, notamment au début du film. Il est donc le parfait exemple d'un film dans lequel on a beaucoup de déplacements, où il faut un peu se débrouiller dans cette matière au plateau.

⁵ Synopsis en provenance de la page Allociné du film.

Tout d'abord le choix a très vite été fait de faire vivre pour de vrai cette ambiance pendant le tournage et donc de jouer toutes les séquences avec une figuration très bruyante y compris hors champ, y compris avec des tirs (du moins quand ils sont visibles à l'image, ils ont été réduits le plus possible). Il y avait un gros enjeu de crédibilité de cet affrontement et de besoin d'énergie de la masse, particulièrement lié à une figuration en très grande partie non professionnelle. Il fallait donc créer l'ambiance pour de vrai au plateau pour porter en même temps les acteurs principaux et la figuration. Pour faire en sorte qu'on croit à cette ambiance et que les spectateur·ices soient en immersion dans cette atmosphère.

Pour répondre aux besoins d'enregistrer ce mélange entre de nombreux mouvements de foule et des interactions individuelles, Arnaud Lavaleix, chef opérateur du son, à préparer une configuration lui permettant d'être mobile et de répondre à de nombreuses situations.

Un perchman était avec une perche HF mono en train de suivre la caméra et ce qu'elle filmait, en essayant d'être toujours au plan. Pour capter les mouvements de la foule et la sensation qu'il y avait de l'activité de toute part, il a fait le choix d'avoir deux couples perchés. Un LCR qu'il avait lui-même à la main et un couple ORTF perché par le second assistant son. Le LCR restait assez collé à la face et lui permettait de capter le brouhaha de l'action au plan de manière moins précise mais plus large que ce que proposait la perche. Ça permet à la fois de combler ce qui ne pourrait pas rentrer dans la perche – parce qu'il y a beaucoup d'actions en même temps et la perche ne peut pas tout couvrir parfaitement à elle toute seule – et à la fois d'offrir une première largeur à l'action. Le choix du LCR pour ce couple a été fait afin d'avoir donc cette spatialisation via l'air apportée par le stéréo, le centre quant à lui permet de venir en soutien à la perche notamment quand la perche vient chercher des éléments qui sont moins en avant plan. Ce LCR a pu aussi, à certains moments, venir en remplacement de la perche. C'est notamment le cas dans une séquence de véhicule où un groupe monte à l'arrière d'une camionnette au cours d'un plan séquence. L'espace qui n'est pas filmé dans la camionnette est extrêmement réduit donc Arnaud Lavaleix a dû se mettre aux pieds du chef opérateur et le perchman ne pouvait pas monter avec lui, le LCR remplace alors la perche. Par ailleurs, à ce même moment, la camionnette bouge énormément. Arnaud Lavaleix a donc installé le LCR sur un casque porté sur sa tête de manière à pouvoir se tenir en même temps.



image en provenance du making off d'Athena

L'ORTF quant à lui servait pour tout ce qui passait en off. Il était pensé à la fois pour pouvoir servir dans les arrières d'une diffusion 5.0, mais aussi pour venir récupérer d'autres éléments et capter le plus d'éléments possibles de cette ambiance très fournie. L'intérêt du multicanal réside aussi beaucoup dans les enjeux de réverbération de tous les sons à fort impact (les coups de feu, les feux d'artifice, les tirs de mortier, etc.). Sur ce type de son l'important pour lui c'est d'enregistrer leur réverbération. Le son du coup de feu en lui-même peut être un peu saturé, il y a de grandes chances pour qu'il soit remplacé en postproduction, mais la réverbération de ce coup de feu dans l'espace est moins facile à recréer. Sur ce type de sons il peut par exemple détimbrer complètement le micro sur des coups de feu pour réduire le niveau de l'impact mais être bien sur l'acoustique du lieu. Ce couple était également en liaison HF pour permettre une grande mobilité. Il y a des moments au cours desquels, notamment lors des déplacements en véhicules, le second quittait l'action pour rejoindre au plus vite le prochain décor et pour être présent sur d'autres parties du plan séquence, la liaison HF paraît donc indispensable. D'autant qu'avec un si grand nombre de personnes présentes et de mouvements, la liaison câblé peut être dangereuse.

En plus de ces micros perchés, il a eu recours à de nombreux appoints. C'est à nouveau le cas dans des séquences véhiculées où des CCM étaient cachés dans les bouches d'aération à l'arrière de la camionnette. Ils servaient à couvrir les petits moments où les acteurs rentrent avant que l'équipe son n'ait la possibilité de venir, ou sur des plans où on filme la camionnette dans une autre voiture depuis l'extérieur mais on aperçoit tout de même ce qui se passe à l'intérieur. Des appoints sont également utilisés sur des intérieurs quand trop d'actions se déroulent dans des espaces séparés.

Quelques micro cravates ont aussi été utilisées, mais positionnées sur des personnages clefs parce qu'il n'y avait pas plus de huit micro cravates positionnés. Pour autant, les costumes des rôles principaux étant très bruyants (kway remonté jusqu'au cou, blousons en cuir, etc.) les HF frottaient beaucoup et l'essentiel du signal était donc récupéré par la perche sur ces personnages. La difficulté résidait aussi dans des enjeux de production : lorsqu'on équipe un figurant au son c'est qu'il dit un dialogue important qui doit être capté. Un figurant avec un micro HF passe alors de figurant à silhouette, avec un salaire plus élevé. Au vu du budget du film (35 millions), il était possible de faire passer quelques figurants en silhouette mais de manière assez situationnelle, ce qui réduit donc le nombre de micros que l'on peut placer. « Je n'aime pas beaucoup l'utilisation des HF sur ce film parce que comme on est très en mouvement je ne peux pas les mixer. Ce n'est pas la première de mes préoccupations. La première de mes préoccupations, c'est que ça rentre bien dans les micros principaux et les micros que l'on a pu cacher. »⁶

Les micros cravates ont parfois été utilisées pour certaines répliques pour anticiper l'arrivée de la perche sur des grosses distances, souvent en complément d'appoints. Dans une telle densité d'actions la perche ne peut physiquement pas être partout et les micros cravates permettent de compenser en partie cet enjeu.

Arnaud Lavaleix a également eu beaucoup recours à des sons seuls pour combler ces manques. Il a principalement pu faire des sons seuls de ponctuels et non des grosses ambiances de foule parce que créer à nouveau la même foule que pendant les prises était impossible en pratique sur le tournage. Il a donc davantage ciblé les éléments à enregistrer séparément. Il s'est presque uniquement concentré sur les sons seuls de voix et de répliques qui occupent une place très majoritaire dans le son du film.

⁶ Entretien avec Arnaud Lavaleix

La gestion de gain représente aussi un enjeu primordial sur ces séquences. On y retrouve des dynamiques très importantes. Elle l'est d'autant plus dans le travail d'Arnaud Lavaleix sur ce film qui suit l'action et est donc toujours avec l'enregistreur dans une version portable (en harnais) et souvent avec le LCR dans une main. Et ce sont les nombreuses répétitions présentes dans ce film qui lui permettent de faire des bons réglages malgré ces contraintes. Pour ce film il y avait trois semaines de répétitions en conditions réelles, donc il a pu faire des premiers réglages de gain assez précis. Notamment sur les HF, il détermine un niveau et ne bouge pas le gain des HF en cours de prise. Ça permet de réduire drastiquement les sources à gérer. Parfois, il regroupe aussi toutes ses sources sur un fader qu'il module alors très lentement. Ça permet de gérer plus rapidement des enjeux de fort niveau quand il y a beaucoup de micros à gérer et une modification significative du niveau global de la séquence.

Les nombreuses répétitions ont aussi permis de supprimer les tops lancés pour le jeu. Les répétitions commencent avec de nombreuses indications pour les acteur·ices et au fur et à mesure, en lien avec la mise en scène, le chef opérateur du son a réussi à faire supprimer les différents tops sonores sans gêner pour autant son bon déroulement. La séquence fonctionnait alors de manière chorégraphique.

En complément des sons seuls au tournage fait par le chef opérateur du son, l'équipe de postproduction son – Matthieu Autin, monteur son et Marco Casanova, mixeur – sont retournés sur le décor avec une équipe de 50-60 figurants rémunérés et le réalisateur pour enregistrer des sons seuls de foule sur deux journées. Avant de mettre en place cette session d'enregistrement, Matthieu Autin a d'abord travaillé avec uniquement les sons directs pendant plusieurs semaines. Il fait un bout à bout sur une session et écoute tous les sons seuls et les ambiances raccord ce qui lui permet de se créer une sonothèque du film. Il commence ensuite le montage de direct et travaille avec la matière qu'il a en allant jusqu'à une étape assez poussée avant d'aller enregistrer ces sons seuls. Ainsi, il a pu éprouver la matière déjà présente et à fait une liste exhaustive de ce dont il avait besoin, en rentrant dans le détail : "un homme qui pousse un cadi en criant pendant 20 secondes, une discussion entre deux ados pendant 10 secondes, un off de policiers qui discutent sur 2 minutes...". Cela lui permet aussi de ressortir des éléments qui étaient dans le direct mais noyés dans la masse de l'ambiance, donc si on a envie de les travailler davantage ça peut être bloquant. L'idée de ré-enregistrer quelques éléments permet donc de regagner cette flexibilité en postproduction tout en gardant une énergie

importante au tournage. Par ailleurs, cette méthode semble plus efficace : au tournage, il est parfois difficile de savoir quel son seul de quelle prise faire quand on vient de tourner énormément de choses très différentes. On a certes une équipe de figuration conséquente et les lieux avec nous, mais sur des séquences importantes comme celle-ci les sons seuls demandent du temps et on bloque donc une équipe à plusieurs reprises ce qui représente un coût important en cumulé. En postproduction c'est très concret, on sait quelle prise a été montée, on enlève toutes les parties qui ont pu être coupées. Et le coût peut être prévu en avance et plus facilement maîtrisé. Tous ces ponctuels ainsi récupérés lui permettent de donner du relief aux séquences et de l'immersion pour le public.

Les couples, il s'en sert surtout pour donner l'impression de masse, pour ajouter à cet effet de densité. Il peut aussi s'en servir pour ramener de l'acoustique sur certains tirs par exemple.

Pour le travail de la version internationale (VI), étant donné que Athena est un film Netflix, les exigences sont très importantes, comme on va le voir dans la prochaine partie – par exemple sur des séquences plus calmes, il vient découper chaque respiration pour les mettre dans une piste option. Il met donc en place trois niveaux d'options pour les séquences de foule. Ainsi, dans l'option 3 il met les foules denses, un peu continues, où on peut avoir des éclats de voix en français si on les écoute individuellement mais qui ne s'entendent pas dans la masse. Ou alors sur un cri sans mot, sans enjeux de langage. En option 2, il met les sons mono qui servent de ponctuels. Ça représente les reliefs, les petites voix avec un peu de langage. Ce niveau d'option peut être utilisé ou non en fonction du nombre d'éléments doublés dans la nouvelle version, du temps disponible pour les faire. Elle peut également être utilisée en guide pour les post synchros d'ambiance. Le premier niveau d'option correspond au niveau plus classique d'option. Ce sont les respirations sonores, les efforts de personnages, etc.

Ils ont écouté les versions allemande et anglaise du film pour voir ce qui était monté, et ils ont remarqué que le niveau trois est énormément utilisé et que le niveau deux peut aussi revenir.

« On dit souvent qu'un film peut se faire avec un seul micro. Celui-là, je pense vraiment pas, sans le multicanal je pense que je n'aurais pas pu le faire. »⁷. Ce film est donc un

⁷ Entretien avec Arnaud Lavaleix

bon exemple que l'évolution des techniques offre de nouvelles possibilités de fonctionnement de tournage concernant la gestion de l'environnement bruyant.

On peut ainsi voir avec cet exemple un cas concret d'adaptation de méthodes de travail aux contraintes d'un film qui peut donner des idées de fonctionnements.

II - INTÉRÊTS ET INCONVÉNIENTS DE CE TYPE D'ENVIRONNEMENT

Avant tout, il s'agit de se demander si ce film ou cette séquence devrait être tourné avec un environnement sonore dense, cohérent avec l'ambiance qu'on cherche dans la séquence une fois le film fini, ou mieux vaut-il le diminuer pour préserver la prise de voix ? La réponse dépend donc évidemment beaucoup des films et des envies de mise en scène. « Il y a des films qui appellent à avoir du bruit, à être un peu plus chaotiques. C'était le cas de *Polisse* de Maïwenn par exemple. Et il y a d'autres films comme ceux de Dupontel où il veut entendre ses acteur·ices chuchoter alors il tourne beaucoup en studio dans le calme » nous indique Rym Debbarh Mounir. Elle précise que la question repose énormément sur le point de vue de la séquence. « Est-ce que le personnage subit son environnement sonore, ou au contraire est-ce que la séquence se déroule dans le point de vue sonore du personnage. ». Dans le premier cas on va plus souvent venir chercher à avoir l'environnement autour du personnage au son et dans le second il est plus courant que nous cherchions à avoir des éléments moins distincts, qui ressortent moins.

On peut donc d'ores et déjà considérer que ce type de dispositif ne s'applique pas à toutes les esthétiques de films même si la palette peut être assez large. Des films travaillant des aspects de déréalisation – comme ça peut être le cas dans des séquences de *Birdman* par exemple – ou ayant des gros besoins de marge de manœuvre et de réinvention en postproduction paraissent ainsi moins propices à avoir un environnement très bruyant dès le tournage. On a alors des éléments avec des dialogues très liés aux ambiances qu'il est très difficile voir impossible – hors post-synchro – de supprimer. Une séquence dans laquelle on voudrait faire évoluer l'ambiance de manière significative au cours de la séquence pourrait sembler difficilement réalisable si elle est tournée dans un milieu bruyant. C'est un exemple qui s'est concrétisé dans mon film de fin d'étude. Sur une séquence montrant l'inconfort d'un des personnages, nous voulions faire disparaître progressivement l'ambiance de la fête foraine pour souligner le malaise. Pour nous permettre de mettre en place cet effet, nous avons tourné cette séquence dans une fête foraine à l'arrêt avec de la figuration silencieuse.

Nous allons ainsi chercher, au cours de cette partie, quels peuvent être les avantages et les inconvénients de tourner avec un environnement bruyant dans le direct afin de nous aider à réfléchir au besoin de chaque séquence, de chaque film.

A - POUR LE JEU

Lors du tournage de mon film de fin d'étude, nous avons donc tourné certaines séquences avant que la fête foraine n'ouvre et les avons terminées quand les manèges autour commençaient à se lancer, que les gens rentraient dans la fête foraine et que l'environnement devenait donc vraiment bruyant. Et c'était très étonnant de voir à quel point le jeu changeait ; même si on s'y attendait, on a pu constater une nette différence. Tout d'abord un changement technique de niveau de voix qui était évidemment plus poussé pour passer au-dessus du niveau ambiant. Mais aussi dans les intentions de jeu, on arrivait à gagner beaucoup de quotidien dans le jeu, les pauses avaient un vrai sens et on croyait simplement beaucoup plus à la séquence. Côté comédien·nes, le ressenti est largement partagé et on me parle de simplicité de jeu, d'évidence.

« L'acteur doit prendre en charge les choses qui ne sont pas présentes au plateau, ce n'est pas impossible ça fait partie de notre formation. Mais ça peut aussi décaler le jeu, le rendre moins intuitif, moins sensible. [Avec du bruit] l'atmosphère change, notre rapport aussi, tout devient plus simple. Ça change le rythme aussi. [...] Le plus flagrant c'est les silences, on est chargés de tout ce qui se passe autour et qu'on n'a pas à imaginer, et ces silences passent d'une pause un peu plus technique à un vrai temps ressenti. Ça change un peu tout. »⁸.

On l'observe aussi sur des choses plus simples du jeu comme les regards qui devaient être un peu dans le vague dans une séquence, évidemment quand la vie autour existe pour de vrai, que l'environnement est davantage réaliste et que des événements sonores apparaissent, on sait où regarder et les positions sont plus évidentes à trouver.

De plus la question se pose plus largement sur les bruits qu'on interdit aux acteur·ices de faire au plateau pour ne pas parasiter la prise de son voix, on interdit aussi parfois certaines actions qui pourraient devenir gênantes par dessus la voix. Rym Debbarih-Mounir nous en parle « Si tu bloques tout, d'un seul coup ils ne peuvent plus bouger et ils vont trop penser à ce qu'il faut faire ou ne pas faire, et ils ne sont pas libres. Ça donne un jeu qui n'est pas du tout intuitif, qui n'est pas libre. Surtout aujourd'hui avec les outils [de denoise] qu'on a. Je préfère qu'on les laisse faire leurs actions, que ça crée de la vie et que le jeu soit juste et nous on atténue les bruits

⁸ Échange avec Guillaume Schmitt-Bailer, acteur dans mon film de fin d'étude

parasites s'ils ne marchent pas, plutôt qu'on ait une belle prise de voix et qu'on soit moins juste au jeu. »

On observe donc que les enjeux de réflexion de ce mémoire valent pour les environnements ayant une nature sonore très dense, mais qu'on retrouve des questionnements similaires dans des séquences plus calmes avec des plus petits évènements. Les possibilités de suppression de ces bruits via les outils de denoise sont plus faciles par rapport aux milieux sonores plus denses. Nos habitudes de réduction de ces bruits systématiques durant le tournage, au même titre que le réflexe systématique de réduction des bruits de bouche en post-production, sont aussi peut-être à remettre en question en fonction des contextes. Ces plus petits éléments paraissent également influencer le jeu, et étant facilement modulable en post-production, Il paraît donc pertinent de se poser la question en lien avec l'équipe de post-production, éventuellement avec la mise en scène également, de savoir quelle place on leur donne au tournage ou à quel point il faudra chercher à les feutrer, à les atténuer. Cette question s'est notamment posée à François Morel dans une séquence de bureau dans *Département 137* où les claviers autour étaient bruyants mais jouaient dans l'ambiance de la séquence et du plateau. Le choix a été fait, en lien avec la post-production, de les garder au tournage. Ce contexte a aidé le jeu et n'a pas été un problème dans le travail en postproduction, nombre de ces éléments ont d'ailleurs été gardés dans la version finale du film.

Par ailleurs, les retours de Rym Debbah-Mounir mais aussi ceux de Valérie Deloof⁹ et Julien Sicart¹⁰ nous montrent qu'il y a donc une plus value certaine du côté du jeu naturaliste lorsqu'on tourne dans un environnement bruyant correspondant à la séquence. Il semble que cette vraisemblance du quotidien dans ce type de milieu sera toujours plus crédible et nous fera croire davantage à la séquence en tournant réellement dans un environnement bruyant. Et peut-être que c'est ce petit truc en plus, cette vraisemblance, qui fait qu'on croit réellement à la séquence, qu'on s'y projette, qu'on dépasse la fiction pour entrer dans l'émotion.

« Si on va tourner dans une église on va avoir de la reverb d'église, et c'est souhaitable d'en avoir pour que ça soit cohérent, pour que ça fonctionne au jeu. Et bien de la même

⁹ « Il est clair que ce qui est autour des acteur·ice influence leur jeu, leur façon de bouger, de réagir, leur manière de parler. La place que ça a dans la mise en scène paraît évidente » *Entretien avec Valérie Deloof*

¹⁰ « Je l'ai remarqué plusieurs fois, si tu les mets dans l'ambiance [les acteur·ices] ça vas vite sonner beaucoup plus juste, on va y croire plus facilement » *Entretien avec Julien Sicart*

manière si on va tourner une séquence en milieu bruyant il faut qu'il y ait du bruit. Plus il y a l'environnement sonore autour, plus les acteurs jouent avec et donc plus t'es dedans en tant que spectateur. »¹¹

Cependant, si la recherche de la séquence n'est pas une séquence naturaliste, si par exemple on veut avoir des personnages qui chuchotent entre eux dans une forme d'intimité au milieu d'un carnaval, alors tourner dans un endroit réellement bruyant semble contre-productifs : les acteur-ices ont de grandes chances d'avoir du mal à s'entendre tout d'abord et il faudra enlever tout le bruit voir faire de la post synchro pour rentrer dans leur intimité. On va donc devoir à l'inverse tricher cette intimité et perdre ce naturel qu'on cherchait à gagner, les acteur-ices se forçant à être dans une situation difficile à trouver en milieu bruyant.

B - POUR LA MISE EN SCÈNE

Côté mise en scène, tourner en milieu bruyant permet de se projeter davantage : plus on a un résultat proche du final, plus il est simple d'appréhender si tels ou tels enjeux de mise en scène fonctionnent. C'est un mouvement qu'on observe globalement dans nos métiers, on le retrouve par exemple en image avec de plus en plus d'aboutissement dans les notions de pré-étalonnage des films en amont et sur le plateau, ou de pré-montage durant le tournage. Aller dans ce sens en aidant les metteur·euses en scène à être plus proche de ce que sera la séquence finale paraît donc être pertinent et correspondre à ce mouvement. Cet enjeu peut avoir de l'influence à plusieurs endroits, outre le simple fait de savoir davantage à quoi va ressembler la scène, ça permet aussi de la ressentir davantage, et par là de ressentir le rythme de la séquence. Dana Farzanehpour, chef opérateur du son, appuie cette idée : « Le réalisateur y croit plus [quand on fait vivre le plateau], ça change vraiment des choses au plateau. Il a plus l'impression d'y être, il est moins angoissé. Donc certainement, il fera peut-être moins de plans ou en tout cas ça lui permettra de mieux se projeter. »

Cela permet certes de se projeter davantage, mais si nous enregistrons ce qui se passe autour de l'action cela veut dire que cette matière doit être intéressante pour ne pas perturber l'action. Particulièrement si c'est de la figuration, il faudra les diriger

¹¹ Entretien avec Rym Debbarh-Mounir

peut-être davantage. Leur donner des thématiques à aborder, des mots clefs à évoquer sans que pour autant tout le monde ne dise la même chose. Pour autant ce travail est souvent effectué ou par les chef.fes de file ou par les assistant.es à la mise en scène qui sont assez habitué.es à diriger de la figuration et c'est un enjeu souvent facilement réalisable.

Au-delà de ces enjeux très concrets pour la mise en scène, la construction d'un environnement bruyant a de l'influence sur différents points qui sont liés aux enjeux de mise en scène. Nous l'avons vu sur les questions de jeu et nous allons continuer à le voir sur les questions de montage.

C - POUR LE MONTAGE

Le son a une place essentielle dans le rythme d'un montage. Une séquence passe beaucoup plus vite quand on rajoute l'ambiance dessus et c'est le son qui a tendance à rythmer une séquence – comme nous le montre notamment Daniel Deshays¹². C'est un constat qui vaut peu importe la séquence mais qui est d'autant plus vrai que le niveau de l'ambiance est important. Ainsi c'est un point central pour les séquences se déroulant dans un espace dont l'identité sonore est censée être riche.

Avoir cette ambiance directement présente dans les rushs permet de construire un rythme bien plus juste dès les premières versions. Les silences ne sont plus des vides un peu flous mais semblent regagner en sens. On se rapproche d'un rythme organique, d'un mouvement plus naturel que ce que l'on peut construire à partir d'une image plus silencieuse ou du moins sans environnement fourni. Ce sont autant de points d'appui pour structurer un montage.

Pour autant une des limites dans ce type de dispositif se retrouve dans la possibilité d'évolution au montage. Et c'est peut-être une des limites les plus importantes et qui se lie de près aux enjeux de prise de risque qui va avec l'enregistrement d'une ambiance sonore dense en même temps que les dialogues. Une séquence pensée comme étant une séquence très dense placée à un moment du film où il est intéressant d'être dans une immersion avec ce qui entoure les personnages pourrait être placé autre part dans le film

¹² « Dans l'énergétique du film, c'est surtout le son qui fait travailler notre sensibilité à la vitesse et au flux. » Daniel Deshays, *Entendre le cinéma*

et devenir une séquence où des plans ne seraient pas montés et dans laquelle on voudrait davantage être dans l'intériorité des personnages en limitant les interventions extérieurs. Même si dans tous les cas la présence à l'image de l'environnement incite généralement la présence des éléments bruyants, des évolutions d'intention de séquence au montage sont fréquentes et un changement drastique de volonté de construction sonore peut survenir. Il revient ainsi aux chef·fes opérateur·ices du son de penser leur prise de son de manière à faciliter ces évolutions. On peut par exemple envisager une ambiance raccord détimbrée pour servir d'ambiance plus lointaine. Ou des appoints placés de manière à laisser des possibilités de variations de points d'écoute. Cependant, même en offrant davantage de possibilités, ces options restent très ancrées dans une mise en scène avec une présence importante de l'ambiance autour de l'action. Sur une volonté de déréalisation d'une séquence par exemple, comme on l'évoquait en introduction de cette partie, les possibilités restent très limitées et la prise de son avec des directs denses pourrait contraindre cette option.

La prise de son au sein d'un environnement silencieux permet beaucoup plus de flexibilité à ce niveau. En n'imposant pas une ambiance sonore sur les dialogues, on retrouve davantage de marge de manœuvre en postproduction sur le type d'ambiance qu'on veut construire, son niveau, etc. Ainsi, les enjeux d'effets et de traitement que nous voudrions apporter uniquement sur la voix sur des séquences déréalisés seraient plus difficilement applicables sur les voix. On peut par exemple imaginer une séquence se déroulant dans un rêve où on voudrait facilement ajouter de la réverbération sur les voix des personnages principaux de manière différente que sur l'environnement qui entoure les personnages. Si cet axe vient après au montage, le tournage en milieu bruyant rendrait contre-productif cet aspect du travail sur cette séquence. Ce choix doit donc se faire dans un échange avec la mise en scène afin de prendre une décision éclairée.

On peut pour autant nuancer ce propos en rappelant qu'au même titre que des retake sont envisagés en cas de changement important au montage, des séances de post synchro sont aussi organisées en cas de modifications de répliques, de problèmes de jeu ou de problèmes techniques. Elles peuvent donc également être envisagées en cas d'évolution drastique des intentions sonores. L'enregistrement du dialogue – sans l'environnement bruyant autour – en son seul au tournage est aussi une sécurité qui peut être prise. Ce n'est donc pas non plus un choix sans aucun filet qui pourrait venir mettre en péril une séquence entière en cas d'évolution puisque des solutions sont possibles.

Tourner dans un environnement sonore dense implique également d'ajouter des accidents dans le direct. Ce point en particulier peut être un avantage comme un inconvénient, si un gros incident arrive au tournage ça peut être un élément qui réduit le choix des prises en rendant une prise inutilisable. C'est un exemple que j'ai pu expérimenter dans le cadre de mon film de fin d'étude où un manège de fête foraine avait remis une musique plus forte que sur la prise précédente, ou dans une autre prise où une boucle de son d'un autre manège s'était lancé au mauvais moment et venait trop perturber l'attention et était difficile à supprimer sous les dialogues. Nous avons ainsi dû écarter ces prises. Dans ces cas particuliers on remarque que la musique est un facteur pouvant être particulièrement problématique car difficilement nettoyable. Il y a donc un risque que quelques accidents viennent rendre la prise plus difficilement exploitable – même si du nettoyage en postproduction son ou de la post-synchronisation reste envisageable. Pour autant certains accidents peuvent aussi donner une tension à une séquence, donner de nouvelles idées en montage. Flavia Cordey, cheffe opératrice du son et monteuse son, nous dit « Tant que le son qu'il y a autour est cohérent avec la narration alors il peut exister. Même un avion qui passerait au-dessus du tournage, s'il est cohérent avec l'environnement qu'on construit dans la séquence – en échange avec la mise en scène – ça peut créer des choses très intéressantes même émotionnellement dans la séquence donc c'est pas nécessairement à jeter. »

On pourrait penser que les coupes sont alors particulièrement contraintes par le son mais en milieu bruyant, particulièrement s'il est dense, la masse de bruit va être facilement travaillable, coupable, modulable. C'est ce que nous avons pu expérimenter dans le cadre de mon film de fin d'étude où même sur des parties du montage avec un rythme de changement de plans relativement dynamique, les coupes n'ont jamais été dictées par l'environnement sonore autour. Alors même que les plans avaient été tournés à plusieurs heures d'écart au milieu d'une fête foraine avec beaucoup de monde. Mais le fond évoluait assez peu et restait plausible dans une densité de fête foraine. Pour autant le raccord entre les fonds peut être plus problématique si le niveau de l'environnement est moins important. On a par exemple eu plus de soucis avec un champ contrechamp tourné avant l'ouverture de la fête foraine : un plan tourné dans le silence et un autre tourné avec une soufflerie permettant aux ballons de tir à la carabine de voler. Le travail des raccords pour permettre les coupes a ainsi été plus

problématique pour nous sur une séquence “classique” et moins sur une séquence tournée dans un contexte particulièrement bruyant. Mais si la quantité de figurants varie d’un plan à un autre, alors le niveau de fond peut varier et ainsi complexifier les raccords et donc demander très tôt dans le montage une attention au raccords son.

Enfin, cela suppose un dialogue plus fort entre montage image et montage son. Savoir ce qui peut facilement être enlevé, nettoyé, voir si tel élément est un problème ou non peut être difficile à déterminer si on n’utilise pas fréquemment les outils de denoise. L’envoi, en cours de montage, de certaines séquences peut permettre ainsi d’anticiper les problèmes et de permettre une suite de postproduction plus fluide. Il pourrait ainsi être envisageable d’imaginer une première projection de détection – projection au cours de laquelle on regarde ce qui doit être post-synchronisé – faite plus tôt, en cours de montage, potentiellement avec une équipe plus réduite.

D - POUR LE TRAVAIL DU SON

Sur certaines séquences, le travail du montage son peut être facilité. En effet, si on tourne avec un environnement muet alors que la séquence se déroule dans un univers bruyant, il faudra reconstruire les différentes couches de sons parfois de manière laborieuse. « Si on a une figuration qui fait semblant de parler en arrière-plan, elle va exagérer automatiquement les mouvements parce que ça n’est pas naturel du tout de faire semblant de parler. Et ensuite, au montage son, on va avoir du mal à mettre une ambiance qui fonctionne parce que même si l’ingé·e son prend une ambiance super avec la figuration, souvent elle ne va pas coller à l’image, on ne va pas y croire parce que la figuration n’a pas joué de manière naturelle dans le plan. »¹³

Sur certaines séquences ça peut donc être plus facile du point de vue de la postproduction d’avoir l’ambiance enregistrée dans le direct. « Tout est là déjà, donc je n’ai presque rien à faire, c’est là et ça fonctionne. Je nettoie les quelques éléments qui ressortent, j’ajoute quelques sons ponctuels mais c’est tout. »¹⁴. On l’a vu, les outils

¹³ Entretien avec Rym Debbarh Mounir, monteuse son

¹⁴ Entretien avec Valérie Deloof, monteuse son, sur une séquence dans une fil d’attente de boîte de nuit dans *Enzo* de Laurent Cantet et réalisé par Robin Campillo

actuels de réduction de bruit permettent une finesse de traitement qui rend ce type de travail plus rapide et plus efficace qu'auparavant.

Pour autant, cette présence d'une base solide dans les directs nécessite tout de même souvent l'ajout de matière en montage son en prévision du mixage. En effet, en ayant des directs très chargés, on réduit aussi la marge de manœuvre au mixage. Sur les enjeux de traitement d'une part : si on applique un traitement sur la perche on va nécessairement l'appliquer sur les voix des acteur·ices mais aussi sur l'ambiance autour. D'autre part, sans ajout de sons supplémentaires la possibilité de travail sur la spatialisation de l'environnement est assez réduite. Il est donc important de travailler des ponctuels au montage son et, si un traitement particulier paraît nécessaire, ça permet également d'isoler certaines répliques du reste de l'ambiance pour retrouver une partie de cette flexibilité au mixage.

Évidemment, cette efficacité n'est pas automatique. Elle suppose bien sûr une prise de son suffisamment maîtrisée en amont, une bonne gestion des rapports signal sur bruit et un environnement en cohérence avec les besoins narratifs. Par ailleurs, dans des séquences très denses en mise en scène, comme c'est le cas dans des films comme *Athena*, un travail plus conséquent paraît inévitable. Mais quand ces conditions sont réunies, le direct devient une matière première déjà proche du résultat final. On revient à avoir moins d'ajouts, moins de construction et probablement plus d'ajustements ou de précisions.

Une autre particularité concrète du tournage en milieu bruyant se retrouve de manière assez évidente dans le niveau de bruit présent au plateau. Ce constat, assez anodin, peut impliquer beaucoup de contraintes. Tout d'abord c'est une contrainte pour la·e chef·fe opérateur·ice du son qui, dans certains forts niveaux, va avoir beaucoup de mal à entendre son mix, même avec des casques assez isolants. Il faut alors se fier au vu-mètre et ré-écouter les prises. Le problème est le même à la perche, dans une grande densité il peut très vite être compliqué d'entendre les dialogues et le suivi du dialogue change et nécessite de se caler sur les expressions du corps. Ce n'est pour autant pas toujours facile étant donné qu'une forte densité de bruit peut aussi être accompagnée d'une forte densité de personnes ou d'obstacles autour des comédien·nes, ce qui peut rendre compliqué le suivi précis. Mais ces forts niveaux sont un enjeu à prendre en compte pour l'ensemble de l'équipe et du tournage. Une forte densité sonore en continu

peut être très fatigante. D'autant plus qu'avec une forte densité de personnes il faut faire attention au matériel, circuler est moins évident, communiquer ensemble peut aussi être compliqué si le bruit ne se coupe pas entre les prises. C'est un effort supplémentaire important qui influence toute l'équipe et qui doit donc être pris en compte dans la conception du temps de travail. Cet enjeu est d'autant plus important pour les tournages se déroulant en décors en activité, n'étant pas contrôlés par le tournage, puisque le bruit ne s'arrête pas entre les prises.

E - POUR LA PRODUCTION

À première vue, tourner dans un environnement bruyant peut sembler plus lourd pour la production. On doit organiser l'environnement sonore durant la prise pour avoir quelque chose de cohérent, ça paraît – par manque d'habitude – moins maîtrisable et donc une perte de temps potentielle dans la journée, il faut par ailleurs enregistrer tout de même des sons seuls et il reste du travail au montage son. Tout dépend pour autant de la manière dont le projet est anticipé.

Si les ambiances sont utilisables et que les voix sont travaillées dans un bon rapport signal sur bruit, il y a moins de sons à refaire, moins de travail de reconstruction. On peut avancer plus vite, parfois même éviter certaines séances de post-synchronisation de foule par exemple.

Évidemment, il y a des cas où ce type de tournage demande plus de moyens. Organiser des sons seuls avec des figurants après le tournage, faire revenir une équipe sur le décor, embaucher des personnes en plus pour ces sessions ça représente un coût. Mais ces dépenses doivent être concentrées sur quelques journées ciblées, et elles permettent d'obtenir une matière sonore précise, adaptée et sur mesure si faite après un premier travail de montage son. Cette méthode paraît donc plus efficace que de chercher à recréer une ambiance en postproduction. Elles semblent par ailleurs prévisibles et maîtrisables. Dans certains cas, ces sons participent fondamentalement à l'ambiance du film, le coût semble relativement moindre au regard du coût total.

Ce qu'il faut retenir, c'est que ce type de tournage n'est pas forcément plus coûteux ou plus lent, mais il demande d'être bien pensé. Il faut que la production sache à quoi s'attendre et que l'équipe soit alignée sur les objectifs de mise en scène sonores. Dans certains cas et notamment sur des séquences avec une mise en scène moins mouvementée, ça peut même être plus simple et au final plus économique.

III - PARTICULARITÉ DES SONS RYTHMIQUES

Nous nous sommes jusqu'à maintenant intéressé·es aux environnements bruyants d'une manière générale en traversant les manières de répondre aux spécificités de ces environnements ainsi qu'en nous questionnant sur leurs avantages et leurs inconvénients. Pour autant nous avons exclu de notre champ de recherche les musiques et plus largement les sons rythmiques qui représentent des particularités notables de fonctionnement.

Au cours de cette partie, nous allons donc nous attarder sur les environnements sonores incluant des sons rythmiques. Est compris la musique dans un premier lieu, élément rythmique le plus courant, mais aussi tout élément ayant un rythme propre, répétitif et identifiable qui impose une continuité sonore. On peut par exemple inclure les machines d'une usine ou des ventilateurs industriels ou même certains sons de mer qui peuvent avoir un rythme répétitif et identifiable.

A - CONTRAINTES EN JEUX

En décors maîtrisés

Prenons par exemple le cas classique d'une séquence dans une boîte de nuit, se pose systématiquement la question : comment joue-t-on la musique au plateau ? Dans ce type de séquence on retrouve beaucoup d'enjeux, de méthodes de travail dont le questionnement et la manière de faire est un questionnement souvent incontournable avec la mise en scène. Ces scènes demandent une énergie particulière. La question de comment diffuser la musique, quelle organisation trouver en fonction des besoins de la séquence est donc une question incontournable dont la charge revient principalement à l'a·u chef·fe opérateur·rice son.

Si la musique est absente, ou jouée à un niveau trop bas pour préserver la prise de son, le niveau de voix des acteur·ices sera probablement trop bas. S'il y a de la danse, le rythme et l'énergie de celle-ci peuvent être facilement aléatoires. Les comédien·nes semblent déconnecté·es, et en étant en dessous du niveau il peut être difficile d'inclure ensuite la musique sans que la séquence ne paraisse irréaliste. On est alors dans un entre-deux fragile, le jeu peut paraître peu crédible et déconnecté du contexte et de l'environnement. Clémence Peloso en parle dans son mémoire "justesse des voix des

acteur·ice·s”¹⁵ : « Jouer avec un niveau fort de voix sans entendre la musique rendra la justesse sur le plateau étrange. Les acteurs ne maîtrisent pas ce qui va être fait car il y a un effet de post-production, ça les fragilise. ». Cet enjeu de justesse des acteur·ice est donc un enjeu majeur que l’on retrouve particulièrement dans ce type de séquences et probablement même davantage que dans les autres types d’environnement sonore que nous avons traversés jusqu’à maintenant. À l’inverse, lorsqu’on envoie la musique au niveau auquel elle est censée être réellement durant la séquence, l’énergie se met en place immédiatement : le corps suit, la voix se pose. Cela permet également d’avoir des moments où naturellement les acteur·ices parleront moins fort parce qu’on est dans un petit creux de musique comme des moments où il faut parler beaucoup plus fort, voire où iels ne s’entendent presque plus, parce qu’on est dans un pic de niveau, alors iels se rapprochent l’un de l’autre et hausse la voix pour s’entendre. Ça crée du jeu et du réalisme à la séquence, ça la crédibilise du point de vue du spectateur également. Pour autant cette méthode implique beaucoup de contraintes.

Dès qu’il y a l’enjeu de passer d’un plan à un autre la coupe se retrouve alors particulièrement contrainte par la musique : si elle ne se fait pas dans un raccord parfait avec la musique diffusée la coupe devra soit être très assumée soit se sentir énormément. C’est donc un choix esthétique important, dicté par la musique et non pas par le jeu ou par les actions de la séquence. Certains films utilisent cependant cette option de diffuser la musique à niveau tout le long de la prise, mais dans des conditions particulières. On retrouve par exemple cette situation dans l’épisode 2 de la série *La Fièvre* (réalisé par Ziad Doueiri) dans lequel le personnage de Sam Berger chante dans sa voiture *Dans les yeux d’Émilie* de Joe Dassin. Le rythme du montage doit nécessairement se baser sur la musique puisque le personnage principal chante en direct. Pierre Gauthier, chef opérateur du son, a donc décidé de diffuser réellement la musique dans la voiture pendant la séquence.

De plus, si la musique ou la version de la musique change en cours de postproduction – ce qui arrive de manière très fréquente notamment pour des raisons de droit – alors la prise de son est inexploitable. Par ailleurs, la musique ne peut évidemment plus bouger une fois la séquence tournée, ni sa version, ni le choix de la partie de musique utile. Ces contraintes étant trop fortes, l’option de diffuser réellement la musique sur le plateau à

¹⁵ Clémence Pelos, *La justesse des voix des acteur·ice·s*, p.29

fort niveau n'est choisie que très rarement, excepté dans des cas particuliers comme celui de *La Fièvre* ou dans des plans séquence.

D'autres méthodes existent cependant. La plus courante est de diffuser la musique pendant les répétitions pour habituer les acteur·ices au niveau auquel iels doivent parler, puis de la diffuser à nouveau avant les prises voir à quelques moments durant la prise pour redonner une référence de niveau et de rythme s'il y a de la danse. Il faut alors être particulièrement vigilant·e à ce que le niveau de voix reste bien soutenu tout au long de la prise, et d'une prise à l'autre. Il est très fréquent si ce n'est pas quasiment systématique que ce niveau de voix chute, il est en effet très peu naturel de soutenir sa voix dans un environnement silencieux. Pour permettre à la mise en scène de se projeter davantage, on peut aussi envoyer au tournage une musique témoin dans les écoutes de la mise en scène, qui peut être déjà pré-traitée, de manière à se rapprocher d'un rendu final et de faciliter la direction d'acteur.

Une autre méthode, de plus en plus courante, est de diffuser uniquement les basses de la musique pour donner le rythme, ces basses pouvant être filtrés ensuite pour retrouver une prise de son sans musique. Ce fonctionnement ne permet pas de diffuser la musique très fort puisqu'on diffuse exclusivement les sons graves. En les diffusant trop fort, ils feront entrer en résonance du mobilier de la pièce ce qui créera des résonances dans des fréquences plus hautes et il ne sera plus possible de simplement filtrer les basses fréquences. Par ailleurs, diffuser uniquement des basses fréquences très fortes ne permet pas de donner un niveau de référence crédible aux acteur·ices pour leur permettre de jouer facilement avec le bon niveau de voix. Cependant, il permet de rappeler continuellement la présence de la musique et donc qu'il faut soutenir le niveau de la voix tout en donnant un rythme aux acteur·ices ainsi qu'aux figurant·es pour danser sur un rythme similaire et d'être plus "dedans" que sans musique. Nous testerons les possibilités techniques et ses limites dans la prochaine partie.

Il est encore possible de travailler avec des oreillettes placées sur les acteur·ices. Cette méthode permet de donner un rythme quand il y a besoin de danser ou d'interagir avec la musique, pour chanter par exemple. Pour autant ça ne fonctionne pas très bien pour donner une référence de niveau de voix auquel parler, ça peut même les déconcentrer. Certain·es chef·fe opérateur·ices du son utilisent aussi d'autres méthodes comme mettre des bouchons d'oreilles aux acteur·ices pour qu'iels s'entendent moins bien et

lèvent ainsi naturellement la voix. Il faut évidemment que la coiffure ou l'axe des plans le permette sans que ça soit visible.

Ces différentes solutions sont mises en place pour permettre d'avoir des danses en cohérence avec la musique pour tenter d'avoir une séquence plus crédible qui se tient davantage. Cependant, ça implique donc d'avoir un choix fait en amont sur la musique de la séquence ou au moins sur le tempo de cette musique pour que les mouvements soient crédibles avec la musique ajoutée en postproduction. Et c'est d'autant plus vrai si les personnages interagissent directement avec la musique, que ça soit du chant ou simplement des réactions, dans ce cas il paraît essentiel d'avoir clearé les droits en amont du tournage non seulement de la musique mais de la version de la musique, ce qui dans le même temps paraît évident et est rarement mis en pratique – hors comédie musicale. Si le choix de ne diffuser la musique qu'au début de la prise est fait, on se retrouve avec des tentatives de suivi de tempo par les acteur·ices et la figuration mais qui sont donc un peu plus flous, plus la prise avance plus c'est approximatif, ça implique donc une moins bonne cohérence entre l'image et la musique mais ça permet également un filet de sécurité en cas de changement de musique.

Dans toutes ces méthodes, c'est une tension permanente entre préserver l'authenticité du jeu ou la qualité du son. Il est aussi possible de considérer que l'énergie à l'image est trop importante, que c'est essentiel d'avoir de la musique à fort niveau comme c'est le cas pour des séquences de *Climax* réalisé par Gaspar Noé¹⁶. Il faut alors faire un son dit témoin au tournage et enregistrer ensuite les dialogues en post-synchro. Aujourd'hui de nouvelles options sont possibles avec le développement du denoise par IA. Nous allons l'expérimenter dans la prochaine partie.

Ces enjeux sont tout de même à relativiser d'après Ken Yasumoto « dans beaucoup de séquences de boîte de nuit on peut avoir une partie de danse mais avec rarement un dialogue important à l'intérieur, puis les personnages vont aller un peu à part de la fête pour avoir un échange significatif dans la narration. Donc quand on est vraiment dans la fête on peut diffuser la musique et par exemple la couper dès que le dialogue commence. Une fois que les personnages se sont éloignés la musique est censée être moins forte et donc l'enjeu de devoir parler au dessus du niveau de la musique est

¹⁶ Entretien avec Ken Yasumoto qui a fait le son de *Climax* réalisé par Gaspar Noé

moindre. » Donc dans ce contexte particulier les contraintes de mise en scène permettent de s'en sortir sans trop de difficulté.

Ainsi, on retrouve donc deux enjeux principaux dans ce type de séquence, la crédibilité du jeu (avec une attention particulière portée sur le niveau de voix) et pour les séquences musicales le rythme de danse.

En décors contraints

On l'a vu dans la partie précédente, la continuité rythmique impose des contraintes fortes de post production et de mise en scène si elle est jouée en direct sur le plateau. Ces contraintes sont d'autant plus importantes si la présence de ces sons rythmiques est une contrainte imposée par le lieu, et particulièrement dans le cas de la musique (fête foraine ou carnaval par exemple).

Pour éviter ces éléments très contraignants de montage, il faut accepter de nouvelles contraintes au tournage. Prenons ainsi le cas d'une séquence tournée dans une fête foraine en activité. Dans ce type de dispositif les éléments musicaux ne sont pas contrôlés par l'équipe de tournage mais posent pour autant les mêmes problématiques. On se retrouve avec les mêmes contraintes de montage et les limites des raccords obligatoirement très francs. Il faut donc trouver des solutions, négocier un niveau de diffusion de la musique et jouer avec l'effet de masque en partant du principe que l'environnement bruyant au-delà de la musique – présent au tournage et ajouté en postproduction – masquera la musique ou du moins suffisamment pour faire passer les coupes.

Il s'agira par exemple de penser le plan de travail de manière à éviter ou réduire la présence de ces sons rythmiques. Ça a été le cas lors de mon film de fin d'étude. Un stand diffusait de la musique à très fort niveau proche d'un de nos lieux de tournage. Nous avons alors dû construire le plan de travail de manière à tourner les plans dans lesquels on voyait moins l'arrière-plan avant que ce stand n'ouvre et nous avions une petite fenêtre de temps de tournage où le stand fonctionnait sans musique le temps de le mettre en route avec déjà du monde dans la fête foraine pour tourner les contrechamp dans lesquels nous avons besoin de vie en arrière plan. Nous avons aussi dû adapter le découpage, nous avons resserré certaines valeurs de plans et ajouté un peu de figuration pour nous permettre de tourner avec le stand fermé et nous avons tourné un plan à deux

de dos plutôt que de face pour tourner quand la musique était lancée. Le son direct était alors un son témoin et nous avons utilisé les doubles des plans précédents pour recréer le dialogue, ce qui est évidemment beaucoup plus facile quand les personnages sont de dos et qu'on ne voit pas bien leurs bouches, nous n'avions en effet pas la possibilité d'organiser de séance de post synchro. Ce plan est au final assez court dans le montage mais nous permet tout de même d'introduire la séquence et surtout l'environnement vivant et bruyant de la fête foraine en montrant une fête foraine active autour des personnages.

La présence de musique en décors contraints peut donc être une problématique prenant une place importante au tournage qui s'avère beaucoup plus bloquante que les environnements bruyants rythmiques que nous avons évoqué au cours des premières parties de ce mémoire.

B - EXPÉRIMENTATIONS

Tentons d'expérimenter les différentes options de diffusion de musique pour discerner les points de vigilance et les limites afin de faciliter le choix de la meilleure option en fonction de la séquence. J'ai donc mis en place des tests en diffusant de la musique à différents niveaux, en essayant de filtrer à différentes fréquences, en essayant de parler par dessus la musique et en essayant les possibilités avec les outils fonctionnants par IA.

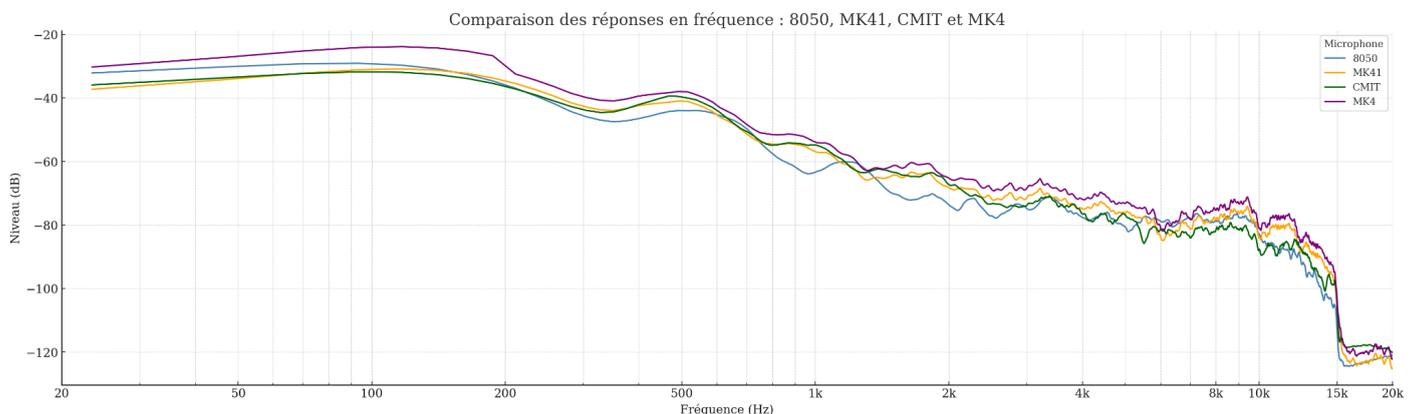
Les tests se sont déroulés avec la diffusion de trois musiques différentes. Les trois musiques sont dansantes, mais avec plus ou moins de présence basses. *Bad Guy* de Billie Elish, une musique avec beaucoup de basses fréquences a tendance électro. *Twist And Shout* des Beatles, avec très peu de basses fréquences. Et *Don't start now* de Dua Lipa, qui est dans un entre deux, davantage pop.

Il faut cependant garder en tête que ces expérimentations ont simplement pour objectif de s'intéresser à chaque méthode dans leurs globalités. Elles permettent ainsi d'en déceler les avantages et les limites sans pour autant pouvoir servir de base certaine dans tout contexte. Je n'ai par exemple fait de tests que dans une seule pièce, et qu'avec trois musiques.

Microphones et directivités

Dans le cadre de ce test j'ai utilisé quatre microphones différents. Deux micros de la marque Schoeps : un CMC6 avec capsule MK41 (supercardioïde), un CMIT (super cardioïde semi-canon). Et un autre micro supercardioïde de la marque Sennheiser. J'ai fait le choix de cette directivité car elle est quasiment omniprésente sur les plateaux pour les séquences dialoguées. J'ai aussi fait des mesures avec un CMC6 et une capsule MK4 (cardioïde) pour voir si une différence surprenante apparaissait. Les micros étaient positionnés non pas avec des capsules coïncidentes, mais en positionnant l'avant du micro sur une même ligne. L'objectif étant de comparer dans des conditions pratiques du tournage avec un micro se positionnant à la limite du cadre. Les micros étaient positionnés à une distance approximative classique de plan moyen.

Regardons d'abord s'il existe une différence de réaction des microphones face à la musique, notamment dans les enjeux de directivité.



On remarque une différence avec la capsule cardioïde, récupérant globalement un peu plus de niveau de la musique. C'est un constat cohérent étant donné la directivité plus large de cette capsule.

On remarque aussi très peu de différence dans les basses et moyennes fréquences de la courbe de réponse en fréquence du CMIT, microphone semi-canon. Le tube à interférence a donc une influence négligeable sur ces fréquences, ce qui est peu surprenant au vu de sa construction. Pour autant on peut noter une diminution des fréquences plus hautes avec, par exemple, 5 dB de moins à 10 kHz que le CMC6 avec capsule MK41.

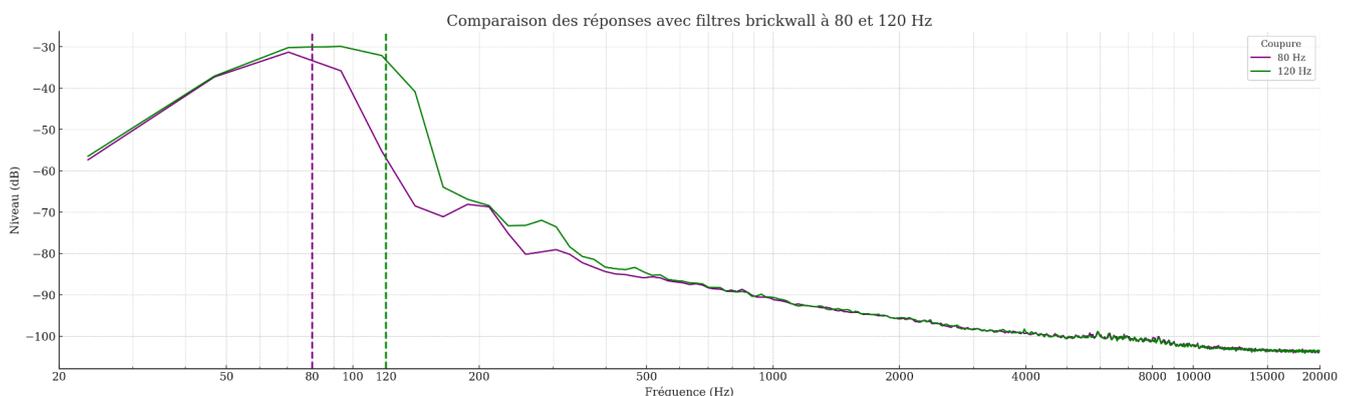
On remarque également une différence entre le micro sennheiser et les micros schoeps avec davantage de basses fréquences, ce qui est à nouveau cohérent par rapport aux caractéristiques du micro. J'ai cependant réduit le gain de ce microphone pour le comparer plus facilement aux autres, celui-ci offrant un niveau de sortie plus important. On peut donc considérer que l'utilisation de microphones supercardioïdes reste pertinente dans ces conditions. Les semi-cansons peuvent également être utiles pour une diminution des hautes fréquences de la musique, donc si elle est diffusée pleine bande.

Diffusion de basses fréquences uniquement

Dans la partie précédente, nous avons également évoqué la possibilité de diffusion de musique avec uniquement les basses fréquences. Comme nous l'avons vu, cette méthode impose un niveau de diffusion inférieur au niveau d'écoute de la musique pleine bande.

Pour mettre en place cette diffusion basse fréquence, on applique sur la musique un filtre passe bas pour lequel on détermine une fréquence de coupure, fréquence au-dessus de laquelle on choisit d'atténuer le signal progressivement en fonction de la pente choisie. Dans ce contexte, il semble pertinent d'utiliser des filtres avec une pente en brickwall pour maîtriser davantage ce paramètre, c'est-à-dire des filtres pour lesquels la fréquence de coupure correspond à la fréquence au-dessus de laquelle on coupe le signal, sans que ça ne soit progressif.

Pour autant, on remarque expérimentalement une différence entre la fréquence de coupure¹⁷ appliquée sur la musique et celle mesurée après diffusion sur les micros de tournage.



¹⁷ J'ai effectué ces tests en mettant appliquant des filtres sur Fabfilter Pro Q4, avec une pente en brickwall

On remarque ici que pour des filtres appliqués sur la musique avec une fréquence de coupure à 80 Hz et 120 Hz, on ne mesure après diffusion une diminution de plus de 3dB qu'à 100 Hz et 140 Hz. Cette diminution se fait par ailleurs de manière progressive, et non pas de manière nette comme ce qu'on attend avec un filtre en brickwall. Cet écart a l'air d'être toujours autour des 20 Hz sur les trois enceintes dont je me suis servi pour ces mesures. Il s'agit donc d'être vigilant·e à cette différence dans le choix de cette fréquence de coupure.

Pour autant, pour supprimer la musique diffusée en postproduction on va à l'inverse utiliser un filtre passe haut pour enlever les basses fréquences et ne garder que celles au-dessus, dans lesquelles se trouvent notamment le signal de la voix. Choisir une fréquence de coupure trop élevée signifie aussi enlever des fréquences de la voix. Les basses fréquences de la voix, particulièrement la fréquence fondamentale et ses premières harmoniques, participent fortement à la perception de certaines caractéristiques du timbre vocal, comme la chaleur ou la présence. Le choix de la fréquence de coupure est donc un choix primordial, entre transmission de la sensation qui donne l'envie de danser et préservation de la voix. Ce choix dépend également des voix des acteur·ices. Ainsi ce choix est aussi influencé par le type de séquence, l'émotion qu'on veut transmettre et le niveau de musique prévu dans le film fini. En effet, dans une séquence de fête avec des personnages qui crient au milieu de la foule et de la musique très fort, avoir une fréquence de coupure plus élevée paraît facilement possible. En effet, en criant les voix deviennent plus aiguës et la musique ajoutée à fort niveau recouvrira probablement les plus basses fréquences par effet de masque. On se rendra peu compte de la suppression des basses fréquences dans les directs et ça gênera moins les spectateur·ices.

Si à l'inverse, nous sommes dans une séquence dans laquelle les personnages parlent légèrement à l'écart de la fête et que la musique est diffusée principalement pour donner un petit rythme aux acteur·ices et à la figuration et que la conversation prend davantage de place, plus intime, alors la perte ou la diminution de ces fréquences fondamentales peut devenir plus gênante.

En fonction des entretiens que j'ai pu réaliser, j'ai eu des avis très tranchés et très différents allant de 60 Hz (pour préserver la voix) à 200 Hz (pour l'aspect fête évoqué). Le choix de cette fréquence de coupure doit donc être réfléchi en fonction de la séquence et devrait se faire, autant que possible, en lien avec la post-production.

Pour autant cette méthode a des limites. Étant donné qu'au-dessus d'un certain niveau – qui dépend de la pièce et des objets qui s'y trouvent – les vibrations dans la pièce peuvent être problématiques pour la suppression de la musique, nous nous retrouvons avec un niveau maximum de diffusion des très basses fréquences. Ainsi, en ayant une fréquence de coupure de notre filtre plus élevée, on peut augmenter aussi le niveau global de diffusion et donc gagner en sensation “d’être dedans”. Par ailleurs, les très basses fréquences donnent surtout le rythme mais ne donnent, la plupart du temps, pas d'autres informations dans la musique. Ken Yasumoto nous fait notamment ce retour dans le cadre du tournage de *Climax* réalisé par Gaspar Noé : « Au début, on cherchait un moyen de pouvoir capter les dialogues en direct, donc on a diffusé uniquement les basses fréquences. Mais assez vite Gaspar est venu me voir et il m'a dit “on va pas y arriver, ils sont pas dedans”. On a donc fait le choix de diffuser la musique à un niveau normal et de post synchroniser les parties dialoguées de cette partie du film. ». Ne diffuser que les très basses fréquences permet donc que les acteur·ices et figurant·es dansent en rythme, mais permet moins “d’être dedans”.

Systèmes de diffusion

Le choix du système de diffusion a lui aussi son importance. Il est influencé par un aspect pratique du tournage – besoin de mobilité – mais aussi par la méthode de diffusion utilisée. Pour de la diffusion des basses fréquences uniquement, il peut être pertinent de passer par l'utilisation de subwoofer. En effet, ce type d'enceinte, spécialisé dans la diffusion de basses fréquences, permettra une plus grande précision dans cette plage de fréquence mais aussi plus de niveau, si l'espace le permet. Cela permet donc de diffuser plus précisément des informations de la musique, notamment si on décide de pitcher la musique avant de la filtrer pour permettre de retrouver davantage d'informations dans la plage de fréquence choisie.

Un autre enjeu de la diffusion de musique se retrouve dans les mouvements faits au plateau : avec les mouvements (de comédien·nes ou de perche), on peut retrouver un niveau variable de la musique ce qui ne facilite pas son nettoyage.

Les variations de niveau induites par le rapprochement de la perche à l'enceinte peuvent être réduites par l'ajout d'enceintes supplémentaires réparties dans l'espace. En diffusant la musique en plusieurs points on diminue le niveau diffusé par chaque enceinte ce qui permet un niveau plus homogène. Cet effet est également renforcé par la

présence de fréquences de résonances. Pour réduire cet enjeu, une égalisation afin de réduire les fréquences de résonance peut être appliquée sur l'envoi de la musique. Le recours à des dispositifs plus lourds comme des line arrays accrochés en hauteur peut aussi être envisagé pour une diffusion plus homogène dans l'espace.

Une autre méthode de diffusion, pour laquelle je n'ai pas pu mettre en place d'expérimentation, pourrait résider dans l'utilisation d'enceintes paramétriques (ou enceintes directionnelles). On retrouve ces enceintes dans les installations en muséographie, dans le cadre de publicités ciblées ou dans certains home-cinémas. Ces enceintes fonctionnent de deux manières. Soit en multipliant le nombre de hauts parleurs de manière à créer des interférences constructives sur la zone dans laquelle on souhaite diffuser la musique et des interférences destructives sur les autres. Soit en diffusant des ultra sons modulés, davantage directifs, qui sont démodulés par l'air (auto-démodulation), ce qui crée de nouvelles ondes dont certaines se situent dans le spectre audible, ainsi l'enceinte ne diffuse que dans un axe précis.

La diffusion de musique via ce type d'enceinte ne permettrait a priori pas la diffusion aux acteur·ices sans que la musique ne soit captée par le micro de la même manière qu'une diffusion classique, mais elle pourrait éventuellement permettre la diffusion de musique à l'équipe de figuration de sorte qu'ils puissent être réellement dans l'ambiance de la soirée et renforcer la crédibilité de cette partie. Il faut cependant être vigilant·es aux résonances sur les murs qui renverraient la musique dans toute la pièce. Pour diminuer cet effet, l'accroche de ces enceintes par le haut – comme c'est souvent le cas en muséographie – pourrait permettre de réduire ce phénomène, possiblement couplé avec la mise en place de moquettes pour absorber le signal, au point que ces résonances soient négligeable ou nettoyable facilement ou recouvertes par effet de masque par la musique ajoutée en postproduction.

Ces différentes configurations dépendent évidemment du temps d'installation disponible et de l'importance de la séquence dans le film, celles-ci pouvant être parfois assez chronophages à mettre en place dans nos dynamiques de tournage.

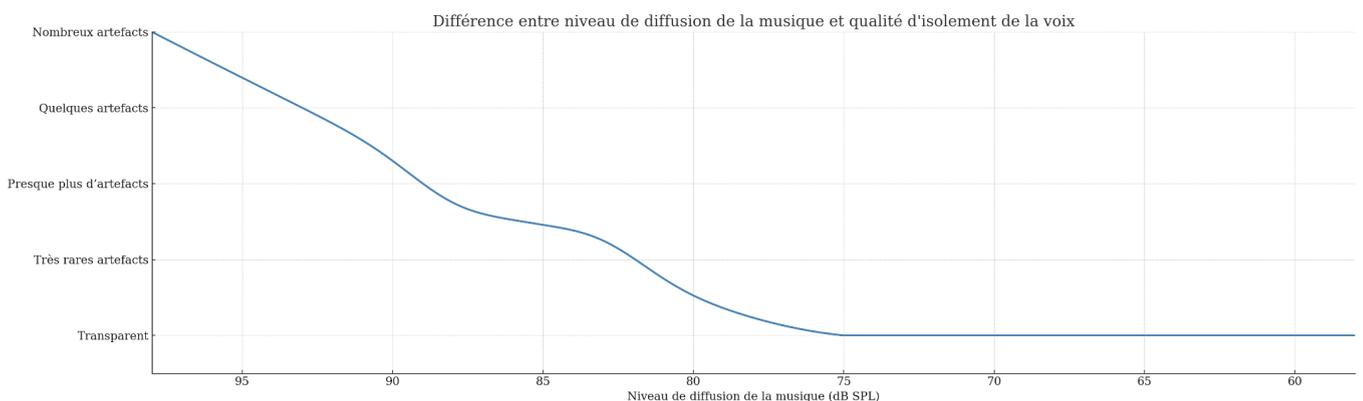
Possibilité de denoise de la musique

Jusqu'à maintenant, la suppression d'une musique des directs diffusée pleine bande était très compliquée. La seule option, outre celle de diffusion des basses fréquences uniquement, était de diffuser la musique à très faible niveau en comptant sur le fait que la musique ajoutée en postproduction cache celle diffusée sur le tournage (effet de masque).

Pour autant, avec l'arrivée des outils de denoise par IA, on a pu voir des outils très performant dans la séparation de la voix et de la musique. Ces outils sont aujourd'hui principalement à destination d'une utilisation purement musicale, pour séparer la voix chantée de la musique (pour faire des versions karaoké ou des reprises par exemple). Mais ils fonctionnent aussi très bien pour séparer une voix parlée d'une musique.

Dans les expérimentations que j'ai pu mener, on peut remarquer qu'il ne semble pas y avoir de différence de qualité de traitement en fonction des fréquences. Une musique avec beaucoup de basses comme *Bad Guy* comme une musique avec davantage de mediums comme *Twist and shout* semblent être interprétés de la même manière par ces outils.

On peut également noter que même à fort niveau ces outils ont une capacité très importante de nettoyage de la musique. Ils ne sont pour autant pas toujours extrêmement constants. On peut observer une augmentation de la qualité de nettoyage sur les trois musiques utilisées à mesure que le niveau de la musique décroît.



Les outils utilisés pour réaliser ces expérimentations sont disponibles en ligne, j'ai fait des tests avec *vocalremover* et *lalal.ai*.

Ces tests ont été faits en considérant un niveau plausible en sensation de diffusion de musique dans une séquence de fête, en étant dans une pièce de taille moyenne. Ce niveau correspondait à un niveau de

référence de 98 dB SPL (mesure avec une pondération C) puis en réduisant le niveau de diffusion par pas de 5 dB jusqu'à 58 dB SPL.

La séparation entre la musique et la voix fonctionne évidemment encore mieux si la musique est moins forte mais que les acteur·ices projettent davantage. De la même manière si les acteur·ices chuchotent, la séparation sera accompagnée de beaucoup plus d'artefacts, et ce même si le niveau de musique est plus faible. Plus la voix est timbrée et plus l'écart entre le niveau de la voix et celui de la musique est important, moins le rendu comportera des artefacts du traitement.

On peut également remarquer que ces outils ne font pas la distinction entre une voix parlée et une voix chantée. Ainsi, dès que la musique contient une voix, il est impossible de la supprimer avec cet outil. Si la musique de la séquence contient de la voix, il faudra donc être vigilant·e à la supprimer pour la diffusion, ce qui peut se faire très facilement avec ces mêmes outils. Se pose cependant la question de la réintégration de la voix dans la version finale du film. Il est ainsi essentiel de préserver cette piste "voix" pour qu'elle soit ajoutée plus tard. Cette séparation n'est, en effet, pas strictement identique à chaque fois que l'opération est faite. Pour retrouver la version d'origine, il faut donc préserver les deux pistes utilisées dans la version diffusée au plateau. Et c'est d'autant plus important si la musique est soumise à des droits musicaux pour que la version finale dans le film ne soit pas différente de la version originale de la musique, clearé.

Pour faciliter son intégration dans le mixage final du film, on peut envisager d'enregistrer en son seul cette piste voix diffusée sur le lieu du tournage (idéalement dans les mêmes conditions de présence de figuration et d'équipe que lors de la prise).

Si les acteur·ices doivent chanter par-dessus la musique, la configuration de diffusion peut être doublée avec des oreillettes pour permettre une diffusion de la voix du chant sans altérer la prise de son direct. Cette méthode peut avoir l'avantage de transmettre l'énergie de la musique (ce qui ne fonctionne pas très bien avec uniquement les oreillettes) tout en permettant une prise de son des voix mais aussi en gardant le guide du chant si cela est nécessaire dans la séquence.

Il est important de noter que ces outils peuvent être très performants avec des voix bien timbrées, donc dans le cas des rôles principaux. Mais dans une séquence de fête avec une équipe de figuration parlante, les voix autour peuvent ne pas toujours être très bien interprétées ce qui peut donner des effets étranges. Ces effets devraient pouvoir être masqués avec l'ajout d'ambiance de foules sans musique, mais dans le cas contraire cela demanderait du nettoyage plus poussé avec des outils de dénoise par algorithme. L'inconstance et le côté parfois un peu imprévisible représente une des grosses limites des outils de dénoise par IA.

On peut ainsi considérer que les possibilités amenées par ces outils sont très intéressantes. En permettant une diffusion pleine bande de la musique à beaucoup plus fort niveau que ce qui était possible auparavant, elle permet de faciliter l'immersion des acteur·ices dans la séquence et de faciliter la crédibilité du jeu, autant dans les corps que dans la voix. Les limites qu'ils ont actuellement nécessitent pour autant de préserver une attention au tournage. Pour autant, au vu de la progression de l'intelligence artificielle, on peut supposer que ces limites vont tendre à se réduire et qu'il pourra peut-être bientôt être possible de diffuser réellement une musique à très fort niveau tout en ayant la possibilité de retrouver des voix de qualité en s'affranchissant des contraintes de montage que la diffusion au plateau de la musique impose.

Se pose également des questions de workflow. Pour que cette méthode permette une liberté de montage, il faut donc que ce traitement soit appliqué avant le montage. Ceci implique donc la création de nouveaux fichiers et par là la suppression des métadonnées permettant la réalisation de l'auto-conformation (étape au cours de laquelle on part du mixdown avec lequel a généralement travaillé le montage pour retrouver toutes les pistes isolées enregistrées au tournage). Il faut donc prévoir une nouvelle étape impliquant d'abord la sauvegarde des métadonnées des mixdowns, puis le traitement de ces fichiers avant le montage avec les outils par IA pour ne préserver que les voix, pour finir par une incrémentation des métadonnées d'origine pour permettre un workflow fluide.

C - SIMILITUDE DES PROBLÉMATIQUES ENTRE LA MUSIQUE ET UN BRUIT RÉPÉTITIF

Certaines contraintes sonores ne relèvent pas de la musique à proprement parler mais produisent des effets comparables en termes de rythme et de découpage. C'est le cas des bruits mécaniques répétitifs : ventilateurs industriels, machines, moteurs, alarmes, cliquetis réguliers, ... Tout élément ayant un rythme propre, répétitif et identifiable qui impose une continuité sonore. Ces sons, bien qu'assez situationnels, finissent par imposer un rythme au montage et à la séquence, au même titre que les séquences avec de la musique diffusée.

Comme la musique, ces sons portent parfois les gestes ou les silences, organisent le temps de la séquence. Le risque, comme avec une musique diffusée en fond, c'est que ce rythme devienne une contrainte rigide à laquelle il faut se plier.

Au montage, ces sons deviennent des lignes rythmiques qu'il faut respecter ou, à défaut, déconstruire proprement. Chaque coupe dans l'image devient aussi une coupe dans un motif sonore et donc un point de fragilité. Couper une alarme régulière ou le martèlement d'un chantier oblige à retrouver une continuité cohérente, sous peine de produire un effet de rupture artificielle. Pour autant ces sons sont plus facilement nettoyables avec les outils de denoise que la musique, on peut donc plus facilement envisager un retravail du rythme avec davantage de souplesse. D'autant que ces sons ont généralement une moins grande notion de continuité que ce qu'on peut retrouver avec la musique. En musique on comprend instinctivement l'évolution de la musique, son schéma narratif. Mais avec cette typologie de sons nous n'avons pas ces enjeux, on ne retrouve pas d'évolution narrative dans le bruit des machines à coudre dans une usine de textile. Par ailleurs, ces sons sont rarement réguliers à la perfection. Une légère variation de tempo, un changement d'intensité ou une instabilité dans le motif peut aisément se justifier. Là où la musique suit une métrique claire, une machine peut accélérer ou ralentir.

Pour avoir cette marge de manœuvre en postproduction, pour pouvoir travailler ces rythmes et faire continuer un son présent dans une prise sur la suivante par exemple, cela impose d'avoir une matière suffisante au tournage. Le plus pertinent semble être l'ambiance raccord (avec l'utilisation d'un couple ou d'un appoint) avec uniquement les

bruits de l'ambiance et sans les dialogues. Mais ces ambiances synchrones peuvent facilement être couplées avec des sons seuls.

Avec ce type de son on se heurte aussi à la question de la redondance : jusqu'à quel point garder le son d'une machine répétitive dans une longue scène ? Comment éviter l'épuisement du spectateur ou l'effet monotone ? Il faut alors trouver des astuces en fonction de la séquence, avoir des variations, fragmenter ou contrebalancer ces sons dans la construction sonore finale, mais cela demande que le montage image soit déjà sensible à ces dynamiques et appui à nouveau les besoins d'échange entre montage image et montage son sur ce type de séquences.

Ainsi, même si ces sons ne sont pas musicaux, ils engendrent des contraintes très similaires dans la fabrication de la séquence. Ils créent une ossature sonore qui impose des contraintes au montage. Et, comme pour une scène musicale, ils posent la question de la cohérence, du tempo, et de la possibilité – ou non – de s'en affranchir.

CONCLUSION

Nous nous sommes penché·es sur les enjeux que représentaient la prise de son en milieu bruyant non seulement pour le tournage mais aussi pour la mise en scène, le jeu et la postproduction. Nous avons abordé différentes manières de répondre aux spécificités des environnements en milieu bruyant en nous questionnant sur leurs avantages et leurs inconvénients. Nous nous sommes également attardés sur le cas particulier des sons rythmiques en s'intéressant plus particulièrement au cas de séquences incluant une diffusion musicale. Tous ces éléments, nous les avons traversés pour tenter de donner des clefs de réponses quant au choix de la densité de l'environnement sonore dans le cadre du travail d'une séquence ou d'un film.

On peut ainsi commencer par considérer que toutes les séquences ne sont pas nécessairement propices à la mise en place d'une ambiance dense sur le plateau, notamment dans les séquences cherchant à ajouter de nombreux effets en postproduction. C'est par exemple le cas des séquences déréalisées, comme les séquences de rêve.

Pour autant, quand la construction d'un environnement bruyant correspond à la séquence, alors cela peut avoir une influence sur de nombreux aspects. À commencer par le jeu pour lequel cohérence entre narration et environnement sonore au plateau semble avoir une influence directe. En ajoutant cette part de réalisme, c'est peut-être ce qui nous fait croire davantage à la situation et permet ainsi de mieux transmettre l'émotion de la séquence. Mais la mise en place d'une atmosphère dense au tournage permet aussi de se projeter davantage dans l'action, que ça soit au tournage ou dès le début du montage. Ces constats valent d'ailleurs autant pour les environnements d'une nature bruyante arythmique comme ceux d'une nature bruyante rythmique.

Ce type de configuration semble ainsi correspondre davantage à des séquences cherchant à construire une esthétique naturaliste, bien que l'influence du milieu bruyant sur le jeu et le travail d'une séquence puisse être également utilisé dans d'autres esthétiques.

La mise en place des milieux bruyants au plateau est notamment facilité par le développement d'outils récents comme le denoise par intelligence artificielle. Que ça

soit dans le cadre de la diffusion de musique comme dans les autres contextes, ces outils permettent un travail au tournage qui favorise les conditions de jeu tout en laissant une marge de manœuvre en postproduction. Ce type de configuration implique cependant une rigueur au tournage que ça soit dans la configuration mise en place, dans la place des sons seuls ou dans la qualité de construction de cet environnement. En effet, la prise de son en environnement bruyant implique une forme de prise de risque au tournage. Celle d'imposer une ambiance au sein du direct ce qui limite les possibilités d'évolution en postproduction. Cela implique de prévoir une prise de son limitant ce risque en facilitant les évolutions, mais aussi d'accepter, parfois, l'arrivée de quelques accidents comme des sons ressortant davantage d'une prise à l'autre. Cette prise de risque implique également le besoin de favoriser des échanges entre l'équipe de tournage et celle de postproduction de manière à ce que les choix faits au tournage n'entraient pas et puissent correspondre à la suite du travail, par exemple sur les questions de format de diffusion.

Nous nous sommes également attardés sur le cas particulier des sons rythmiques en remarquant les contraintes particulières qu'ils impliquent et en étudiant les différentes options possibles en lien avec la diffusion de musique au plateau. On peut ainsi considérer différents outils applicables en fonction des besoins des séquences. En allant de la diffusion de musique à niveau dans le direct – avec prévision ou non de denoïse – jusqu'à la diffusion de basses fréquences, en passant par la diffusion de musique en dehors des dialogues, via des oreillettes, ou pour la figuration grâce à des enceintes directives.

Pour finir, il semble que nous puissions considérer que la mise en place d'une prise de son en milieu bruyant peut apporter une vraie plus-value si elle correspond aux envies de mise en scène, mais elle demande une réflexion des méthodes de travail en amont du tournage en fonction de la séquence ou du film afin de permettre une chaîne de travail de qualité.

REMERCIEMENTS

Je tiens à remercier celles et ceux qui m'ont accompagné·es au cours de ce travail. Jacques Descomps, Pierre Sans, Lucille Waterlot, Olivier Kalonji et Hélène Orfèvre pour vos précieux conseils. Valérie Deloof et Térésa Faucon pour vos retours pertinents sur ce mémoire et votre suivi. L'ensemble des équipes de la Fémis pour leur travail au cours de ces quatre années.

Rym Debbarh-Mounir, Arnaud Lavaleix, Valérie Deloof, Julien Sicart, Flavia Cordey, Matthieu Autin, Dana Farzanehpour, Ken Yasumoto et Lucas Doméjean pour avoir pris le temps de parler de ces sujets dans le cadre d'entretiens. Ils ont porté le travail de ce mémoire et constituent une base indispensable.

Anna Devillaire, Lou Jullien, Juliette Couvreur, Théophile Chauchat et Louis Bellec pour les échanges autour de ces questionnements qui ont permis des prises de reculs essentiels.

Josselin Panchout pour les astuces et partages de réflexions.

Toute l'équipe du film de fin d'étude pour leur investissement, leur motivation et leur bonne humeur dans l'aboutissement de ce projet en lien étroit avec le mémoire.

Juliette Gaudin pour son soutien durant ces quatre années, autant à l'école et dans le travail qu'en amitié

Lana Limar, Louis Bellec, Emilien Pichon et Malo Geffraye-Meriadec, mes camarades de promotion pour les quatre années passées à leurs côtés, les réflexions échangées et le travail commun.

Mes parents, mon frère, ma grand-mère et mes ami·es pour leur soutien continu.

BIBLIOGRAPHIE

Jérémy Marozeau, *L'effet de la fréquence fondamentale sur le timbre*, Université Pierre et Marie Curie - Paris VI, 2004. Français.

Daniel Deshays, *Entendre le cinéma*, collections Klincksieck, 2010 , série « 50 questions »

Clémence Pelos, *La justesse des voix des acteur-ice-s*, mémoire de fin d'étude, la Fémis

Sites internet

Article AFSI « Cannes 2022 - Erwan Kerzanet à propos de "L'ENVOL" de Pietro Marcello », Guilhem Domarcq, 16 mai 2022

[URL:https://www.afsi.eu/articles/97260-cannes-2022-erwan-kerzanet-a-propos-de-l-envol-de-pietro-marcello?utm_source=chatgpt.com]

Netflix studios, Post Production Branded Delivery Specifications, version 5.2

[URL:<https://partnerhelp.netflixstudios.com/hc/en-us/articles/7262346654995-Post-Production-Branded-Delivery-Specifications>]

Amazon studios, Delivery & QC

[URL:https://portal.amazonstudios.com/hc/en-us/articles/15986851525147-Delivery-QC?utm_source=chatgpt.com]

Allociné, page du film *Athena*,

[URL : https://www.allocine.fr/film/fichefilm_gen_cfilm=294589.html]

Entretiens

Entretien avec Matthieu Autin

Entretien avec Flavia Cordey

Entretien avec Rym Debbah-Mounir

Entretien avec Valérie Deloof

Entretien avec Lucas Doméjean

Entretien avec Dana Farzanehpour

Entretien avec Arnaud Lavaleix

Entretien avec Julien Sicart

Entretien avec Ken Yasumoto

FILMOGRAPHIE

Laurent Cantet, Robin Campillo, *Enzo*, 2025

Robin Campillo, *120 battements par minute*, 2017

Ziad Doueiri, *La Fièvre*, 2024, série Canal plus

Romain Gavras, *Athena*, 2022, produit par Netflix France, Iconoclast, Lyly Films

Gauthier Gervaise, *Deux et demi*, 2025

Alejandro González Iñárritu, *Birdman*, 2014

Ladj Ly, *Les Misérables*, 2019

Maïwenn, *Polisse*, 2011

Dominik Moll, *Dossier 137*, 2025

Gaspar Noé, *Climax*, 2018

Céline Sciamma, *Portrait de la jeune fille en feu*, 2019

Nathan Villaneau, *Fort comme un lion*, 2025