



**L'INTEGRATION DE LA POST-
SYNCHRONISATION**

Les pratiques actuelles

Sreesti SHEIKH

Mémoire de Master 2 - Spécialité Son

Directeur de mémoire interne : Alan BLUM

Directeur de mémoire externe : Cyril HOLTZ

Examineur : Sylvain LAMBINET

Référent académique : Laurent MILLOT

Coordinateur des mémoires : Corsin VOGEL

Juin 2023

RESUME

La post-synchronisation est une étape essentielle dans la post-production cinématographique pour assurer une bonne qualité sonore et l'intelligibilité des dialogues. Il s'agit d'enregistrer ou réenregistrer des répliques et certaines présences (notamment des respirations et des réactions) des comédien·ne·s après le tournage d'une production audiovisuelle. En France, cette étape est particulièrement délicate car elle se doit d'atteindre une sensation de naturel dans les dialogues au plus proche du son direct (obtenu au tournage) auquel le cinéma français est très attaché. Sa réussite est liée à la notion de vraisemblance, qui constitue la problématique centrale de l'étude.

Ce mémoire examine les pratiques actuelles, les défis rencontrés et les stratégies mises en place par les professionnel·le·s pour garantir une prise de son de post-synchronisation de qualité ainsi que son intégration aux étapes du montage et du mixage via l'analyse des méthodes et des outils utilisés pour maintenir la continuité et la vraisemblance sonore. Une attention particulière est accordée à de nouveaux outils de mixage spécifiquement développés pour l'intégration de la post-synchro, afin d'observer leur fonctionnement, leurs apports, leurs limites et les besoins auxquelles ils répondent. Cet état de l'art s'appuie en grande partie sur des entretiens avec des professionnel·le·s et des développeurs de logiciels récents pour compléter une documentation peu abondante sur ces pratiques.

Enfin, un test d'écoute est mis en place pour évaluer l'efficacité de ces logiciels, en mettant l'accent sur leur contribution au raccord et à la vraisemblance sonore.

Mots-clés : post-synchronisation, intégration ADR, (Automatic Dialogue Replacement), logiciels, plug-ins, workflow¹, vraisemblance

¹ Dans le contexte de l'industrie audiovisuelle, le workflow fait référence à l'ensemble des étapes nécessaires à la création, à la production et à la post-production d'un contenu audiovisuel, tel qu'un film, une émission de télévision ou un projet multimédia.

ABSTRACT

ADR (Automated Dialogue Replacement) is a crucial step in film post-production to ensure good sound quality and dialogue intelligibility. It involves the recording or re-recording of lines and certain vocal elements (such as breaths and reactions) by actors after the filming of an audiovisual production. In France, this step requires particular attention as it aims to achieve a sense of realism in the dialogue that closely resembles the original on-set sound, which is highly valued in French cinema. The notion of realism and credibility of ADR dialogue is the central issue addressed in this study.

This paper examines current practices, challenges faced, and strategies implemented by professionals to ensure high-quality ADR sound recording and mixing through an analysis of the methods and tools used to maintain sound continuity and realism. Special attention is given to new software and plug-ins specifically developed for ADR mixing in order to observe their functions, contributions, limitations and the needs they address. The information presented in this paper heavily relies on interviews with professionals and recent software developers to supplement the limited existing documentation on these practices.

Lastly, a test is conducted to evaluate the effectiveness of these software tools, with a focus on their contribution to sound synchronization and plausibility.

Key words: ADR-matching, post-synchronization, software, plug-ins, workflow, realism in dialogue

REMERCIEMENTS

Merci à mon directeur interne Alan Blum pour sa patience et son suivi. Merci à Corsin Vogel et Valentin Bauer de m'avoir guidé dans la méthodologie.

Merci à mon directeur externe Cyril Holtz, sans qui ce mémoire n'existerait pas, pour son expertise, sa pertinence, sa bonne volonté et son ouverture d'esprit.

Merci à Florent Fajole pour ses conseils de lecture précieux et sa réactivité.

Merci à Olivier Guillaume, mon parrain professionnel, pour son encouragement, sa bonne humeur, ses conseils et ses corrections.

Merci à toutes les professionnel·le·s et aux équipes de HAL, Izotope et Accentize qui ont pris le temps de participer aux entretiens, m'ont prise en observation sur leurs séances et m'ont confié des extraits de film. Anne Maisonhaute, Antoine Martin, Aymeric Dupas, Catherine Taieb, Cyril Holtz, Frédérique Liebaut, Katia Boutin, Matthieu Tertois, Niels Barletta, Olivier Guillaume et Pierre Picq, vous avez enrichi et nourri ce mémoire, merci pour vos mots et votre gentillesse.

Merci à l'équipe de Polyson pour leur accueil chaleureux pendant tout ce temps.

Merci à toute la promo 2023 pour leur bonne humeur et compagnie rassurante. Merci à Régine Laures, la mère de Nelly Boussellier, de m'avoir donné accès à la suite Microsoft.

Un énorme merci à Elie Bissery qui a relu et corrigé beaucoup de fautes d'orthographe, et qui m'a soutenue tout au long.

Merci à mes ami·e·s de lycée, Amandine, Leela, Riddhidev, Shakthi et Sougandan, pour les appels Discord, les moments partagés au restaurant et pour leur compagnie. Merci à ma famille, qui ne comprend pas un seul mot de ce que je rédige, pour leur soutien à distance. Merci à Dario Benoit-Franco et Krokou Bissery pour leur chaleur et soutien émotionnel.

TABLE DES MATIÈRES

RESUME	1
ABSTRACT.....	2
REMERCIEMENTS	3
TABLE DES MATIÈRES.....	4
INTRODUCTION.....	6
<i>Partie 1 : Pratique de la post-synchronisation, contextualisation</i>	<i>8</i>
I- Histoire de la post-synchro.....	9
a) La naissance des systèmes de post-synchro.....	9
b) Tromper le spectateur·ice : le « raccord parfait », la notion du réalisme	14
II- L'organisation de la post-synchro.....	19
a) Les étapes préparatoires	19
b) L'enregistrement en studio	22
c) Le montage et le mixage.....	26
<i>Partie 2 : L'évolution des habitudes de travail dans la post-synchronisation</i>	<i>27</i>
I- Le choix des entretiens	27
a) Entretiens qualitatifs et semi-dirigés.....	27
b) Le guide d'entretien.....	28
c) Présentation du corpus de professionnels	30
II- Du studio d'enregistrement à l'auditorium de mix	32
a) Le système de travail	32
b) Evolution de la prise de son en studio de post-synchro.....	35
c) Le montage et mixage.....	40
III- Evolution des logiciels récents	44
a) Dialogue Match, Izotope : Un outil global	44
b) Chameleon, Accentize : le <i>reverb-matching</i>	47
c) Boom-It, HAL : L'ajout de mouvement	50
<i>Partie 3 : Les enjeux de la post-synchro, analyse en théorie ancrée</i>	<i>55</i>
I- Le contexte des entretiens	55
II- Principe de l'analyse thématique	55

a)	L'analyse thématique par la théorie ancrée	55
b)	La retranscription	55
III-	Analyse des résultats : Les enjeux de la post-synchros	56
a)	Les catégories	56
b)	La séance d'enregistrement	58
c)	Une post-synchro réussie	66
d)	La vraisemblance recherchée en post-synchro	70
e)	Conclusion	76
	<i>Partie 4 : PPM – Comparaison/Apport des logiciels pour l'intégration des post-synchro</i>	
	77
I-	Exposé du projet.....	77
a)	Présentation générale	77
b)	Choix des extraits	78
c)	Choix et utilisation des logiciels d'intégration	80
II-	Présentation et analyse des résultats	83
a)	Partie 1	83
b)	Partie 2.....	90
III-	Discussions et limites du test.....	93
	CONCLUSION	95
	BIBLIOGRAPHIE.....	97
	ANNEXES	98
I-	FORMULAIRE DE CONSENTEMENT	98
II-	QUESTIONNAIRE PPM	100
III-	Boom-It : entretien avec un des développeurs, Antoine Martin	101
IV-	Dialogue Match : entretien avec les développeurs de chez iZotope	112
V-	Chameleon : entretien avec Benjamin Graf, Accentize	123

INTRODUCTION

La post-synchronisation est une opération qui permet de réenregistrer en studio le dialogue ou la voix off d'une œuvre audiovisuelle dans la même langue et par le même comédien que lors du tournage, pour une raison technique (dans le cas où la prise de son initiale n'est pas exploitable) ou une raison artistique. Nous verrons plus précisément les différents cas qui nécessitent ce travail de post-synchronisation plus tard dans le mémoire. Nous allons employer, au cours de cet écrit, le terme « post-synchro », qui est familier mais plus léger et couramment utilisé par les professionnels.

Les post-synchros étant enregistrées dans un espace-temps différent de celui du son direct, (dans un studio silencieux quelques semaines après le tournage), la question du raccord se pose systématiquement. Ce travail sur le raccord démarre avec la voix de la comédien·ne qui doit se remettre dans les conditions du tournage (portée de voix, timbre, intention du jeu, intonation, etc.). Au-delà du temps passé en studio, un travail d'intégration des post-synchros est effectué par les mixeur·se·s qui cherchent à rendre les voix post-synchronisées indiscernables du son direct.

Ré-acoustiquer les sources, c'est-à-dire traiter les sons pris en studio afin qu'ils retrouvent un timbre et une spatialisation raccord avec les sons du tournage est une problématique à laquelle s'attaquent aujourd'hui des logiciels très performants. Ces logiciels récents ou en cours de développement sont nés pour répondre à certaines demandes dans le contexte de l'intégration des post-synchros aujourd'hui. Ce mémoire essaiera de mieux comprendre ces besoins, en analysant le travail effectué en mixage avec différents outils et différentes tactiques. Il essaiera de mettre en évidence, grâce à l'interaction avec les professionnels, les nouvelles demandes éventuelles qui émergent.

Quels sont les méthodes et outils utilisés par les mixeur·se·s aujourd'hui pour la prise de son et l'intégration des post-synchros dans l'acoustique d'une scène ? Quels sont les apports et limites de ces nouveaux outils dans le cadre du « raccord » et de la vraisemblance ?

Afin de répondre à ces questions, nous nous plongerons dans la littérature existante, allant de mémoires de fin d'études à des essais plus longs et plus précis, ainsi que dans les revues professionnelles. En première partie, nous ferons tout d'abord un passage en revue

historique des techniques d'intégration de post-synchros pour mieux contextualiser le sujet. L'évolution de la prise de son et le mixage de post-synchros fera l'objet de notre deuxième partie, guidée par les entretiens des professionnel·le·s sous forme de citations non-analysées. Nous ferons un point sur les besoins sous-jacents à l'arrivée de nouveaux logiciels sur le marché, et les besoins actuels qui laissent la place à d'autres solutions possibles (technologiques ou organisationnelles). La 3ème partie de ce mémoire consistera en l'analyse de divers entretiens réalisés auprès des professionnel·le·s. Des entretiens qualitatifs et semi-directifs seront menés pour essayer de comprendre les enjeux de la post-synchro. Nous allons également définir « la vraisemblance » dans le cadre des post-synchros grâce aux entretiens. Dans une quatrième partie, nous expliquerons la démarche qui nous a menée à réaliser des intégrations de post-synchros en mixage. Ces extraits destinés à mettre en œuvre les logiciels précédemment étudiés seront ensuite confrontés à des tests de visionnage que nous analyserons pour terminer notre étude. Il s'agira donc notamment d'étudier l'apport des logiciels récents, ainsi que leurs limites.

Partie 1 : Pratique de la post-synchronisation, contextualisation

Il existe deux grandes familles de post-synchro aujourd'hui dans le cinéma français. Les post-synchros d'ambiances servent à construire la matière sonore d'un film. *« C'est toujours une matière hétérogène au film, qui vient nourrir le montage son, en cela, elle est délicate. Elles sont faites sur-mesure pour le film monté, pour qu'elles s'intègrent à la narration, à l'esthétique, au climat. En effet, les ambiances ne « meublent » pas, elles participent à la construction sonore du film, et sont liées au point de vue du mixage. »* (Liebaut, 2019, p.1)²

On fait appel à des comédien·ne·s pour jouer des foules, faire des wallas³ et donner une voix aux figurant·e·s pour enrichir la bande-son. Les voix étant rarement isolées, il n'est pas nécessaire d'avoir les mêmes comédien·ne·s que sur le tournage, ce qui serait difficile à mettre en place logistiquement. Nous ne traiterons pas les post-synchros d'ambiance dans cet écrit car il serait possible d'écrire un mémoire entier dessus.

Il est important de noter à ce stade que nous allons nous concentrer sur le cinéma français de fiction verbo centré. Il s'agit de l'écriture classique du cinéma parlant actuel où la parole est centrale. *« La conception des sons et des images sont consciemment ou inconsciemment orientées dans le but général de faire écouter les dialogues (généralement abondants) comme de l'action, en en faisant le centre –souvent invisible- de jeux de scène, d'effets de montage, d'éclairage, etc., destinés à en faciliter l'écoute et à en dramatiser l'enjeu, tout en effaçant la perception du dialogue comme tel. »* (Chion, 1996, p.94)

Nous allons également restreindre notre étude à des films tournés en prise de vue réelle, excluant ainsi les films d'animation de notre sujet de recherche.

² Il s'agit d'un document non publié rédigé par Frédérique Liebaut, directrice de plateau, s'appelant *Écrit sur les métiers de post-synchro*, 2019. Il ne figurera pas dans notre bibliographie.

³ « Walla » est un terme de l'industrie du cinéma américain, désignant une ambiance de discussions de personnes sans qu'aucune émotion ni mot ne soit perceptible ou intelligible. En France le terme « yaourt » est souvent employé.

I- Histoire de la post-synchro

a) La naissance des systèmes de post-synchro

La post-synchro est née aux États Unis en 1932 à la suite de l'arrivée du cinéma parlant et du son direct ; c'est une conséquence directe de la technologie du son direct sur plateau. La prise de son sur un tournage peut faire face à de nombreux imprévus : sons parasites, jeu non-intelligible, etc. Pour pallier ceci et avoir une bande son continue, agréable et compréhensible, il est parfois nécessaire de réenregistrer les dialogues une fois le montage image terminé. Voici une liste non-exhaustive de raisons pour lesquelles cette étape peut être utile :

Raisons Techniques :

- Bruits parasites ambiants qui nuisent à la compréhension de la parole (cascade d'eau, circulation, club disco bruyant etc.)
- Intelligibilité difficile due à une sous-articulation
- Reconstitution historique (bruit de voitures modernes lors d'un tournage de film d'époque)
- Difficulté technique (distorsion, son sous-modulé, micro-cravate inexploitable etc.)
- Réalisateur·ice souhaitant parler pour donner des indications de jeu pendant une prise

Raisons Artistiques :

- Changement d'intention de jeu d'acteur
- Un ajout de texte ou de dialogue pour un personnage
- Modification du plan sonore (rapprocher ou éloigner spatialement les dialogues)
- Vouloir une plus grande netteté du son (ex. : enregistrement spots publicitaires télévisés)
- Une volonté artistique (on retrouve cette esthétique souvent dans le cinéma italien)

Ce mémoire portera sur le contexte français de la post-synchro, qui évolue dans une « culture du son direct », c'est-à-dire la pratique de prioriser le son direct qui est illustrée par les mots suivants :

« Les sons, les bruits et les voix naturelles, comme elles sont recueillies par le micro, ont un pouvoir de suggestion qu'on n'arrive pas à obtenir avec la post-synchronisation. [...] » (Antonioni, 2003, p.70-71)

Ceci est en effet la posture de plusieurs réalisateur·ices et comédien·ne·s français·e·s, qui résulte dans la réduction de la post-synchronisation en une étape de « rattrapage technique » plutôt qu'une étape créative alors que, comme toutes les étapes de la post-production, les post-synchros ont à la fois un aspect technique et un aspect artistique. *« Pour des raisons historiques et culturelles, elles ont été longtemps considérées l'étape désagréable mais néanmoins nécessaire pour des raisons essentiellement techniques à la qualité du son du film. C'est une étape peu investie artistiquement et surtout vue comme le lieu où la qualité du vivant ressentie dans le direct ne pouvait pas être atteinte et qui plus est, pouvait être perdue. »* (Liebaut, 2019)

Or aujourd'hui, il y a de moins en moins besoin de réaliser des post-synchros suite à des raisons techniques grâce aux logiciels de nettoyage puissants, ainsi que des remplacements possibles par des doubles.⁴ Ceci laisse plus de place pour une créativité qui servirait la narration du film.

Afin de réaliser une séance de post-synchro il existe plusieurs systèmes. Les anglophones utilisent le système de l'ADR (Automated Dialogue Replacement) aussi appelé *looping* ou *looping session*. En France, la culture de la « bande rythmo » (cf. p11) prime. Ces systèmes ont leur racine dans les techniques de doublage développées dès les années 1930.

⁴ Le remplacement d'une syllabe, mot ou phrase par une autre prise, souvent avec un jeu et rythme similaire, qui permet de s'affranchir de l'étape de post-synchro.

- Le « looping »

Le looping consistait à découper un film en courtes séquences. Chacune d'entre elles faisait la durée d'une réplique du personnage, d'une dizaine de secondes. Le morceau de film était monté et joué en boucle, ce qui permettait une projection de la scène répétée autant de fois que souhaité une fois chargée dans le projecteur. Le·la comédien·ne pouvait alors réécouter en boucle, s'imprégner du rythme de la phrase à refaire. Au bout de quelques répétitions suite auxquelles le directeur·ice artistique⁵ lui donnait des indications, iel indiquait à l'opérateur·ice du son qu'iel était prêt à enregistrer. La machine d'enregistrement optique était alors déclenchée et le texte enregistré une ou deux fois.

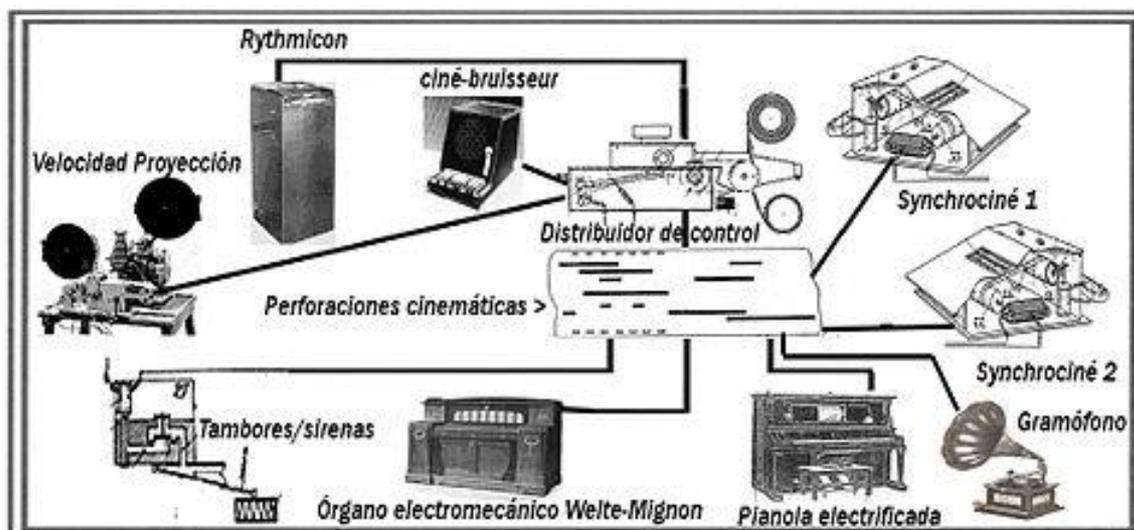
Il n'était pas possible de vérifier le résultat du jeu et du synchronisme avant que le film négatif du son ne soit développé. Cette méthode, utilisée pendant une vingtaine d'années en doublage était critiquée pour donner lieu à des doublages avec une diction mécanique due à la concentration des comédien·ne·s à réciter un texte par cœur et synchrone.

Plusieurs solutions ont été essayées pour permettre aux comédien·ne·s de lire un texte synchronisé avec l'image afin de leur éviter l'apprentissage du texte. Le premier de ces procédés était le Système Topoly, inspiré du dispositif synchro-ciné⁶ de Charles Delacommune, permettant de décaler le déplacement d'une bande par rapport au mouvement d'un appareil cinématographique.

Le texte était écrit à la périphérie de disques verticaux tournant en synchronisme avec l'image à une vitesse de 2 tours par minute. Les comédien·ne·s lisaient le texte écrit d'après les mouvements des lèvres des personnages tandis que défilait devant eux un index situé en haut du disque. Il y avait autant de disques que de rôles à doubler.

⁵ Le·la directeur·ice artistique, aussi appelé directeur·ice de plateau peut donner des indications aux acteur·ices pour faciliter un raccord technique tout en restant dans l'intention artistique du/de la réalisateur·ice

⁶ Synchro-ciné : un dispositif inventé en 1921 par Charles Delacommune avec l'objectif de synchroniser la projection d'un film avec les sons correspondants.



Système de contrôle synchronisé Delacommune à pleine capacité (jusqu'à 9 appareils différents dans la première version du distributeur, 1922, et 12 dans le dernier, 1930). Tous les appareils sont interchangeables (jusqu'à 9/12 pupitres de lecture simultanément, ou 9/12 bruiteurs, ou toute combinaison de ceux-ci). Le projecteur de film ne compte pas (il fonctionne sur une ligne distincte).

Fig 1.1.1, Système de synchronisation Delacommune 1923, Ortiz Morales, CC BY-SA 4.0

- La bande rythmo

Le Système Topoly ne permettait pas de regarder à la fois l'image du film et le texte synchrone. Le comédien ne pouvait alors pas suivre les mouvements des lèvres en même temps que la lecture, et un perfectionnement a été mis en place. « *Les monteurs travaillaient sur une visionneuse où ils pouvaient voir l'image sur un petit écran en même temps qu'ils écoutaient le son de la piste sonore. Sur le mécanisme de la visionneuse était monté un défiler qui permettait, grâce à une boîte de vitesse, d'entraîner un film 35mm transparent à une vitesse 6 fois moindre que celle du film image (depuis 1950 un rapport de 8 est utilisé). Il devenait alors possible de détecter sur ce film les mouvements exacts des lèvres des comédiens* » (Lerouge, 1996, p.120)

Repérer les labiales des lettres P, B et M ainsi que les ouvertures et fermetures de la bouche permettait théoriquement au dialoguiste du doublage d'écrire un texte français avec un synchronisme sans faille. Le texte détecté est écrit en encre de chine sur une nouvelle bande qui est découpée en petites longueurs comme l'image, synchrone avec cette dernière. Cette

bande est projeté en même temps que l'image, sous l'image du film, ce qui permet au comédien d'alterner l'observation du texte et celle de l'image.

Ce procédé de la bande rythmographique (bande rythmo) est toujours employé en France pour la post-synchro. C'est le seul pays jusqu'à aujourd'hui à utiliser ce système. Le texte défilant était autrefois écrit sur une bande polyester perforée (comme la pellicule), mais se présente sous forme de fichier informatique à présent.

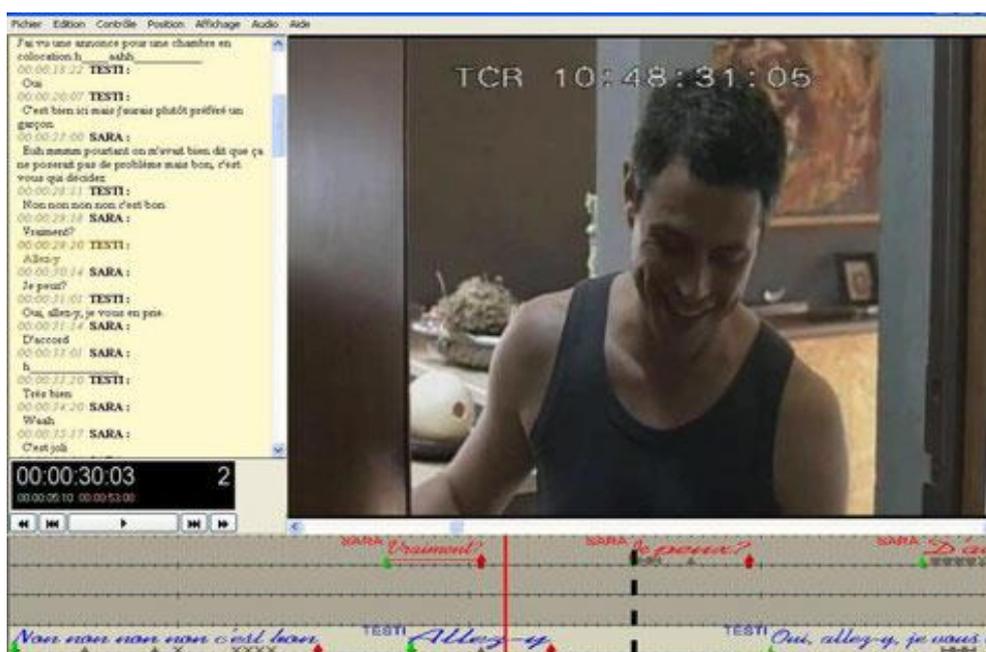


Fig 1.1.2, La Bande Rythmo virtuelle, image prise sur le site d'Antony Penetto
<https://www.anthony-panetto.fr/tradaptation/doublage/technique/>

- L'ADR

Un système très proche du « looping » dans son fonctionnement, l'ADR ou *Automatic Dialogue Replacement*, nécessite un découpage des séquences à post-synchroniser en courtes phrases. L'acteur-ice entend le son témoin à remplacer en boucle, souvent via un casque, et s'imprègne du rythme et de l'intonation de la phrase. L'enregistrement est déclenché au début de la phrase à refaire quand le comédien se sent prêt. « *L'acteur entend, en plus, trois bips sonores dans son casque espacés chacun d'une seconde, et dont le dernier se trouve une seconde avant la première syllabe, pour lui permettre d'attaquer synchrones la première ouverture de bouche* » (Lerouge, 1996, p.122)

Il peut y avoir également un repère visuel de synchronisme fourni à l'acteur·ice, qui consiste en des diagonales de 48 images qui défilent en bas de l'écran (anciennement dessinés sur un film). Le comédien commence à parler quand le trait arrive en bas à droite de l'image. Le repère plus classique reste trois « bips » espacés d'environ 20 images avant le début de la phrase à refaire.



*Fig 1.1.3, photogramme de Postcards from the Edge de Mike Nichols.
Suzanne Vale (Meryl Streep) enregistre des répliques en ADR*

b) Tromper le spectateur·ice : le « raccord parfait », la notion du réalisme

La post-synchronisation est un procédé qui est souvent confondu avec le doublage, les deux ayant pour but de réenregistrer le dialogue d'un film en post-production. La post-synchro, la plupart du temps, sert à remplacer des dialogues sur une scène ou une séquence ; il faut donc pouvoir passer d'une phrase "directe" (enregistrée lors du tournage) à une voix post-synchronisée. On va appeler cela un "raccord voix". Il est donc nécessaire d'avoir le·la même acteur·ice sur l'écran et derrière le micro en studio de post-synchro, cas qui n'est pas appliqué au doublage qui se fait sur toute la longueur d'un film. La post-synchro est également faite dans la langue originelle du film, ce qui ne s'applique pas forcément au doublage.

- Les synchronismes possibles en post-synchro

Il y a plusieurs esthétiques, dues à des synchronismes différents en post-synchro. Le son post-synchronisé domine majoritairement sur le son direct dans les films italiens à partir des années 1940. Esthétique institutionnalisée par les réalisateurs comme Fellini, c'est devenu un pilier de l'étape de la post-production sonore en Italie. Même si le procédé technique porte le même nom, les Italiens cherchaient un synchronisme différent à celui des Français lors des post-synchros.

D'après Michel Chion, ce que les Français considèrent comme un défaut de post-synchronisation dans le son original des films italiens est en réalité un synchronisme plus large et plus ouvert, qui n'est pas réglé au dixième de seconde près. Cette différence est particulièrement perceptible au niveau de la voix : tandis que les synchronismes très serrés s'alignent strictement sur les mouvements des lèvres, les synchronismes plus larges prennent en compte l'ensemble du corps parlant, y compris la gestuelle. En somme, un synchronisme plus large produit un effet moins naturaliste, plus poétique et plus décontracté.

Même s'il nous a semblé important de citer le cas du cinéma italien, nous n'allons pas traiter ce cas particulier car c'est un sujet à part qui a déjà été abordé dans un mémoire. (Shukla, 2022)

La catégorisation de Michel Chion est néanmoins intéressante car cette précision de la post-synchro française justifie des développements technologiques afin d'atteindre un synchronisme « étroit et serré », telle la bande rythmo qui est utilisée de manière exclusive en France. La volonté esthétique qui prédomine la post-synchronisation dans ce pays aujourd'hui, est celle d'un « raccord parfait », un résultat « réaliste », une concordance parfaite entre les mouvements de bouche et la parole. « *La synchrèse, phénomène incontrôlable, amène donc à établir instantanément un rapport étroit d'interdépendance et à rapporter à une cause commune, même s'ils sont de nature et de source complètement différentes, des sons et des images qui n'ont souvent que peu de relation dans la réalité. Le cinéma utilise abondamment cet effet, qui permet notamment la post-synchronisation et le bruitage* » (Chion, 1994, p.433-434).

Daniel Deshays caractérise le synchronisme comme étant un rapprochement entre le monde cinématographique et la réalité, une réduction de leur distance (Deshays, 2010). Sa

notion de fusion, qui établit un lien indissociable entre l'image et le son, est similaire à la synchrèse de Michel Chion. La désynchronisation, quant à elle, devient alors le lieu où les espaces et les temps coexistent sur un même plan.

Mark Kerins présente une perspective plus concrète sur ce sujet. Le·la spectateur·ice, en partie piégé·e par le synchronisme, accepte cette condition (Kerins, 2010). Cela renvoie au contrat audiovisuel : « *La manière avec laquelle la vue et l'ouïe, par leur relation conjuguée suscitent en nous des interprétations narratives et émotionnelles d'une richesse bien supérieure à leur simple addition* » (Chion, 1997, Première partie de l'audio-vision). La bande sonore est constamment exposée au risque de révéler l'hétérogénéité des matériaux et des sens, par exemple à travers un décalage labial en post-synchronisation dans le contexte du cinéma français.

- Le raccord technique

Le son au cinéma est un art de la reconstruction ; il faut raccorder les différents éléments (son direct, synthétisé, post-synchronisé) entre eux afin de créer une cohérence pour que le·la spectateur·ice perçoive un seul espace-temps. La notion de raccord est celle d'une continuité dans notre perception. L'acoustique, la couleur ou timbre de la voix et la projection de la voix sont des éléments auxquels l'oreille est sensible, et un changement soudain à un de ces éléments sans changement perçu à l'image serait déstabilisant pour le spectateur·ice (Dutrieux, 2021). C'est que l'on appellera le raccord technique de la voix.

Du fait de l'importance donnée au synchronisme et au raccord à l'étape de la post-synchronisation, les sons témoins sont d'une grande importance. C'est l'élément qui permet d'avoir une base de travail pour les monteur·se·s, directeur·ice·s de plateau et les comédien·ne·s. Il est possible de post-synchroniser une ou plusieurs scènes, séquences ou la totalité d'un film, même si ce dernier cas est assez rare aujourd'hui. Le raccord technique est établi, la plupart du temps, par rapport au son direct existant.

Nous allons voir plus loin les différents paramètres qui jouent dans ce raccord et comment l'équipe de post-production son parvient à le créer.

- La vraisemblance

Étymologie du mot « vraisemblance » : Formé au XIV^e siècle à partir de l'adjectif vrai (veri) et du substantif semblance (simili), vieux mot signifiant apparence, aspect.⁷ Il existe plusieurs définitions dans le dictionnaire Larousse ou le centre de ressources national textuel et lexical (CNRTL).

« Le caractère de vérité possible de quelque chose »

« La probabilité proche de la certitude »

« Caractère de ce qui semble vrai, juste, aux yeux du sens commun »

La vraisemblance, sans pour autant créer une certitude absolue, donne une apparence de réalisme. Elle est définie par le sens commun et la perception de la réalité. Dans le domaine de l'art, le terme "réalisme" est souvent associé à la représentation fidèle de notre monde, de la nature et de la réalité dans laquelle nous existons. Or, dans le cinéma et par extension dans notre étude, être vraisemblable ne signifie pas nécessairement être réaliste. En d'autres termes, le réalisme cinématographique se rapporte à la question suivante : est-ce que l'œuvre est une représentation de ce qui constitue notre monde ? Le cinéma qualifié de "réaliste" est celui dont l'intention est de refléter une situation réelle.

Il est possible d'aborder la notion de vraisemblance sonore au cinéma en se référant aux travaux de Michel Chion, qui propose deux concepts clés pour la décrire : le son qui sonne vrai et le son qui est vrai. Nous reconnaissons instinctivement une voix qui sonne vrai grâce à notre expérience quotidienne du son. Cependant, au cinéma, il existe d'autres codes et des conventions qui nous aident à évaluer la vraisemblance d'un son. Cela signifie que, dans une salle de cinéma, nous mettons de côté notre expérience sonore quotidienne, afin d'accepter la réalité sonore du monde cinématographique : ce qui sonne vrai.

« Un art d'enregistrement comme le cinéma a créé des codes spécifiques de vérisme, liés à sa propre nature technique » (Chion, 1994 chap. 5).

⁷ Définition selon le *Grand Dictionnaire encyclopédique Larousse* (GDEL), édition 1982.

Un exemple concret de ce phénomène est le rapprochement des voix par rapport au plan de placement des personnages à l'image. Le plan sonore au cinéma ne correspond pas toujours à la profondeur de champ perçue à l'image. Cela ne reflète pas la réalité du monde dans lequel nous vivons, mais dans le contexte cinématographique, il est parfaitement vraisemblable d'entendre les paroles d'un personnage situé loin dans le champ de manière claire et proche. En effet, grâce à l'utilisation de microphones de proximité (tels que les micros-cravates sans fil), nous nous sommes habitués à une écoute des voix plus proche dans le cinéma. Nous constatons que les définitions de la vraisemblance énoncées précédemment doivent être remises en question dans le cadre des arts représentatifs.

Le raccord technique et la vraisemblance ne sont pas des termes interchangeables, mais dans la majorité des cas, l'un ne va pas sans l'autre. Dans le cadre de la post-synchro, la vraisemblance s'approcherait d'un phénomène qui nous permettrait d'accepter une synchronie entre l'image de l'acteur·ice et le son de sa voix. C'est la recherche de cette vraisemblance qui motive une intégration de post-synchro. Nous essayerons de développer davantage cette notion avec l'aide des professionnel·le·s interviewé·e·s plus loin dans ce mémoire.

II- L'organisation de la post-synchro

a) Les étapes préparatoires

- La séance de détection ou de détermination

Pour connaître les répliques à refaire ou à rajouter, la production d'un film organise une projection dite « de détermination » (pour déterminer les post-synchros à faire) ou « de détection » (c'est le fait d'écrire la rythmo). C'est une étape qui se réalise souvent en présence de le·la réalisateur·ice, le·la chargé·e de la post-production, le·la monteur·se image, le·la monteur·se son, le·la monteur·se des directs, le·la mixeur·se ainsi que le·la directeur·ice de plateau de post-synchro (si différent·e de le·la monteur·euse des directs) et le·la recorder⁸ de l'auditorium (Avaress, 2017).

« La détection » est une étape qui sert à isoler les moments où il y a potentiellement besoin de faire de la post-synchro de dialogue ou d'ambiances, les scènes à travailler avec une intention particulière, etc. C'est l'étape fondatrice du travail de post-synchro. Elle intervient la plupart du temps après la sortie d'une version quasi-définitive du montage image et après quelques jours ou semaines de montage des directs - aussi appelé montage parole.

Dans le cas idéal, le·la monteur·se parole aura déjà parcouru le film : iel aura travaillé sur le nettoyage de dialogues et aura cherché des doubles. Ce travail consiste à remplacer une partie de dialogue qui contient un problème technique ou un jeu qui ne convient pas avec une prise synchrone similaire. Ce travail en amont permet de se rendre compte de la matière existante, ce qui facilite le travail de sélection et de priorisation des séquences à refaire en post-synchro.

⁸ Assistant·e technique qui assure le bon déroulé de la séance.

- L'écriture de la bande rythmo

Une bande rythmo est ensuite réalisée le plus souvent avec Mosaic⁹ pour la séance. La bande rythmo numérique est une zone défilante en bas de l'écran sur laquelle des dialogues sont inscrits. Elle défile horizontalement sur la vidéo de manière synchronisée. Les syllabes sont plus ou moins étirées pour correspondre à leur prononciation, de sorte qu'elles passent sous une barre verticale au moment précis où elles doivent être énoncées (cf. fig 1.2). La bande rythmo facilite le travail de le·la comédien·ne qui lit le texte défilant sous la barre rouge de synchronisation.



Fig 1.2.1 La bande rythmo sur Mosaic
Mosaic Logiciel pour le Doublage, Facebook

À la suite de la séance de détermination, le·la directeur·ice de plateau définit également un « cue sheet » pour l'organisation de la session de post-synchro en studio. Cette « cue sheet » indique les répliques à reprendre en post-synchro et la raison de cette reprise. Les séquences et les time-codes sont notés également. On y trouve aussi la catégorie de la post-synchro à faire : personnage principal, secondaire, ambiances etc. C'est le fil conducteur de la séance d'enregistrement de post-synchros (voir image en dessous).

⁹ Un logiciel permettant de calligraphier les répliques et présences (respirations, rires etc.) des personnages de manière synchronisée à l'image.

marie-Line et son Juge

22b 23b 24b 25b 26b 27b 28b 29b 30b 31b 32b 33b

26/02/20

BOUCLE	TIME CODE	CATEGORIE																		
		GENRE	F	H	H	P	P	S	S	S	S	O	O	A	A	A	A			
		Marie-Line	Juge	Père	Alexandre	Louise	Samuel	Tobias	Avocate	Jehanna-	Policier	Policier	Serveuse	Amb	Infirmière	Client	FE 1	HO 1	Patron	FE 2
3 Brasserie	01:01:05:01																			
3 b	01:01:26:21							X												
7A Marché	01:04:30:18	X																		
7 b	01:04:45:13																			
7 c	01:04:57:14	X																		
7 d	01:05:45:19	X																		
9 jardin	01:05:59:13																			
9 b	01:06:34:12																			
9 c	01:07:38:18																			
11 Ext Escaliers	01:08:02:22	X																		
13 Brasserie	01:10:58:06								X						X		X	X		X
13 Amb	01:11:28:17															X	X			
13 b	01:11:43:12																X	X		X
14 Ext Brasserie	01:12:25:04								X											
14 ext	01:13:49:09	X																		
15 Ext	01:14:08:03	X																		
8 Devant Chez Ale	01:15:08:16	X																		(X)
18 b	01:16:30:12	X																		
18 b	01:16:49:17															X				
19 Tribunal	01:17:13:22	X														X				
19 b	01:18:46:23		X																	
19 c	01:19:00:20		X																	
19 d	01:19:13:19		X																	
19 d	01:19:27:07		X													X				
19 e	01:19:48:16		X													X				
20 couloir	01:20:12:16															X				
21 Salon	01:21:49:14																			
23 Brasserie	01:22:50:19															X		X		
24B ext	01:24:20:21	X																		
26 Synagogue	01:26:15:21	X																		
29B Salon Juge	01:28:22:16		X																	
1000	01:30:46:02																			X
2D Voiture Pluie	02:02:21:22		X																	
33B Cuisine	02:02:55:10	X																		

Fig 1.2.2 Cue Sheet de Marie-Line et son Juge, directrice de plateau Anne Maisonhaute

- Les éléments de montage

Du côté du montage des directs et du montage son, il y aura un travail sur les « bouchages » avant la séance d'enregistrement. Ceci consiste à construire une continuité d'ambiance pour avoir un fond sur lequel les post-synchros seront enregistrées. C'est un élément important à avoir au moment de l'enregistrement car il permet de déposer les voix sur un tapis (un fond d'ambiances) et se projeter dans la manière dont la post-synchro s'intègre dans la scène. Il évite d'avoir un « trou » à l'écoute.

Le·la monteur·se des directs organisera les sons suivants par pistes en consultant le·la mixeur·se qui enregistre les post-synchros¹⁰ :

- Pistes perches + HF pour enregistrer et trier les post-synchros
- Un chutier pour garder les pistes non-utilisées
- Ambiances raccords ou bouchages : il est important d'avoir des fonds cohérents.
- Les directs VI (version internationale sans paroles)
- Des pistes FX (effets audio) et musique si elles existent

L'assistant·e image fournit à l'auditorium de post-synchro une version image, éventuellement iel donnera un format différent (plus léger) de la même version de montage au directeur·ice de plateau si demandé. Iel doit parfois également envoyer une version à la personne qui resynchronise et monte les post-synchros si ce n'est pas le·la chef·fe de plateau. On incruste à chaque fois dans l'image le nom de la personne (ou du prestataire) à qui l'image est fourni, ainsi que le Time Code¹¹ (pour que tout le monde se cale de la même façon), et les numéros de séquences.

b) L'enregistrement en studio

- Les conditions de la post-synchro française

Une fois toutes les étapes préparatoires terminées, l'enregistrement de la post-synchro va se faire en studio. En France, les post-synchros se font souvent dans un espace avec un volume important, ce qui permet à le·la comédien·ne d'adopter la posture corporelle qu'iel sent

¹⁰ Informations tirées de *Chaine Post-Prod Son avec Post-synchro*, document rédigé et fourni par Frédérique Liebaut et Anne Maisonhaute

¹¹ Le Time Code ou TC est une référence temporelle utilisée dans le domaine de l'audiovisuel, pour la synchronisation et le marquage de matériaux enregistrés. Le timecode (TC) s'exprime en heures, minutes, secondes et images.

nécessaire, s'écouter et se regarder sur l'écran pour la synchro labiale dans les conditions de salles de cinéma. D'autres pays préfèrent utiliser une cabine séparée pour isoler le·la comédien·ne, équipé·e d'un casque, qui parle dans un micro sur pied en restant plus ou moins statique.

La présence d'un·e perchman ou perchwoman en auditorium en France ainsi que le grand volume de l'auditorium, permettent aux comédien·ne·s de bouger, ainsi que de porter la voix. Cette manière de prendre le son s'applique surtout sur des films de catégorie A. Les films avec moins de budget ou les séries n'ont souvent pas le budget de faire appel à un perchman ou une perchwoman de post-synchro. Dans ce cas, un micro sur pied (souvent un Neumann U87) sera placé devant le·la comédien·ne, et la prise de son sera beaucoup plus statique.

C'est un exercice difficile que de rejouer de manière synchrone en lisant le texte qui défile sous l'écran. La portée de voix dans un auditorium silencieux est également difficile à doser. « *Ce n'est pas facile de revenir faire un rôle des mois après le tournage et beaucoup de comédiens ne réagissent pas bien en regardant leurs scènes pour la première fois [...]* » (Purcell, 2012, p.266). Nous reviendrons sur l'importance du jeu afin d'intégrer les post synchro plus-loin.



Fig 1.2.3, Post-Synchros de Marie-Line et son juge avec Matthieu Tertois (mixeur), Pierre Picq (perchman) et Michel Blanc (acteur)

- Les équipements techniques

Le studio ou l'auditorium de post-synchro est équipé afin de rendre le processus d'enregistrement le plus efficace possible. L'acoustique de ces studios est évidemment de la plus grande importance. Dans les années 40-50 l'acoustique de ces lieux était très « vivante » avec un temps de réverbération entre 0.8-1.5s. La tendance ensuite dans les années 70-90 a été de revêtir les parois du studio de matières absorbantes, pour « assécher » le son afin d'introduire une réverbération de raccord artificielle (Pommier, 1988). Historiquement, les post-synchros s'enregistrent dans des auditoriums de mix ou des studios de bruitage, même si récemment un studio dédié uniquement aux post-synchros a été ouvert chez Polyson Post-Production. Les studios sont souvent équipés de panneaux acoustiques modulables qui permettent de régler la quantité de surface absorbante et réfléchissante. Ceci permet une plus grande variété d'acoustique à la source (la prise de son), ce qui permet la flexibilité de reproduire des espaces plus ou moins grands.

Ces studios sont souvent équipés en LCR (3 haut-parleurs gauche-centre-droite), 5.0 ou en 5.1 pour la diffusion du son. Dans le studio nous retrouvons également :

- Une surface de contrôle et des éléments de traitement de son
- Une station de travail Pro Tools
- Une interface audio pour l'entrée des micros
- Un écran de projection
- Des casques pour l'ingénieur du son, le-la perchman ou perchwoman, l'acteur·ice et les autres personnes présentes si besoin

- Les différents personnes présentes en séance de post-synchro (Pommier, 1988, p.49-56)

Selon l'ampleur du projet, il peut y avoir différentes personnes présentes dans la salle.

➤ L'équipe de réalisation

Le·la réalisateur·ice et parfois des producteur·ices sont présent·e·s pour guider le jeu d'acteur. Iels vont assurer la qualité de l'interprétation verbale et aider à remettre les interprètes dans le contexte et l'émotion du tournage.

➤ Le·la monteur·se image, qui connaît parfaitement le film, les raccords et les répliques.

➤ Le·la mixeur·euse ou ingénieur du son

C'est celui ou celle qui règne sur toute la technique. Son rôle est d'enregistrer les comédien·ne·s en respectant au mieux les conditions de tournage. Iel choisira avec son perchman ou sa perchwoman le type de micro à utiliser (souvent le même modèle que sur le tournage), son placement et son niveau. Iel est également responsable de l'acoustique du lieu d'enregistrement, modulable grâce aux panneaux acoustiques de l'auditorium. Iel passera du temps à ajuster tous ces paramètres pendant les répétitions de chaque scène à post-synchroniser. Iel peut également mettre des effets de filtre ou de réverbération provisoires pour faire écouter le résultat à ceux présent·e·s dans le studio, ce qui permet de mieux rendre compte du résultat final.

➤ Le perchman ou la perchwoman :

Iel se concentre pour être à la bonne distance de la bouche de l'interprète. Iel doit comprendre et reproduire les mouvements de perche nécessaires pour être raccord avec le timbre dans le son direct. Son placement est primordial à la qualité de la post-synchro.

➤ Le·la directeur·ice de plateau :

Iel s'assure de la qualité globale des post-synchros sur un film et est chargé·e de planifier la session de manière efficace pour obtenir les meilleurs résultats dans les délais impartis. Iel fait le pont entre le pôle technique et le pôle réalisation et peut donner des indications aux acteur·ices afin de faciliter un raccord technique tout en restant dans l'intention

artistique du/de la réalisateur.ice. Nous parlerons de l'évolution de ce rôle dans un second temps, car c'est un post qui a connu quelques nouveautés au cours des dernières années.

➤ Le-la recorder de l'auditorium

Iel aide à l'installation et la configuration de l'équipement et assure le déroulement technique de la séance.

c) Le montage et le mixage

Le « recalage » consiste à ajuster temporellement les post-synchronisations afin de respecter la synchronisation labiale des dialogues, ce qui est essentiel pour une intégration réussie. Ce travail minutieux contribue à l'illusion que les paroles des personnages correspondent parfaitement à leurs mouvements labiaux à l'écran. Il est souvent méconnu à quel point le processus de montage des post-synchronisations peut être long et chronophage. Dans le cas de recalage des respirations et présences, chaque son doit être soigneusement ajusté pour correspondre aux expressions et aux actions des acteur.ice.s à l'écran. Cela implique souvent de nombreuses itérations, des ajustements précis et une attention méticuleuse aux détails.

L'étape de montage des post-synchros implique également la sélection et la combinaison des enregistrements de post-synchronisation avec les enregistrements en direct ou d'autres prises de post-synchronisation afin de recréer les répliques nécessaires.

Le mixage est l'étape finale dans la post-production d'un film, et donc de l'intégration des post-synchros. C'est ici que le raccord technique pourra être poussé à son maximum. Nous parlerons plus en détails des méthodes et des logiciels utilisés dans la partie suivante.

Partie 2 : L'évolution des habitudes de travail dans la post-synchronisation

I- Le choix des entretiens

« *Les entretiens exploratoires ont pour fonction de compléter les pistes de travail suggérées par les lectures préalables* » (Blanchet & Gotman, 2007, p.41)

Au cours de nos recherches, nous avons constaté que la bibliographie concernant la post-synchro française, surtout dans les années récentes, était très peu fournie. Afin de poursuivre notre état des lieux de la post-synchronisation à l'heure actuelle et de l'évolution du travail, nous avons donc décidé d'appuyer cette partie sur un retour d'expérience de professionnel·le·s, ainsi que sur des documents de travail internes (non publiés) fournis par certain·e·s d'entre eux. Nous ne manquerons pas de citer les professionnel·le·s quand nos propos s'appuieront sur leur expertise. Une analyse plus approfondie des entretiens, utilisant la méthode de la théorie ancrée sera par ailleurs effectuée plus tard dans cet écrit, en partie 3.

a) Entretiens qualitatifs et semi-dirigés

Au cours de nos premiers échanges avec des professionnel·le·s du domaine de la post-synchronisation, nous avons constaté que chacun·e avait un parcours et une perspective différents en raison des films et des expériences vécues avec chaque équipe. Il était important de recueillir l'expérience unique de chaque personne interviewée avec attention et précision. « *L'entretien s'avère utile lorsque l'enquêteur veut analyser la perception des participants vis-à-vis de leurs pratiques, leurs expériences ou lorsqu'il veut déterminer les valeurs et les normes qu'ils valorisent.* » (Blanchet & Gotman, 2010, p.27)

Il nous a paru pertinent de mener des entretiens qualitatifs et semi-dirigés. Les entretiens qualitatifs servent à « *recueillir des données qualitatives (témoignages, notes de terrain, images vidéo, etc.)* » dans le but de les « *analyser [...] de manière qualitative (c'est-à-dire en extrayant le sens plutôt qu'en les transformant en pourcentages ou en statistiques)* » (Paillé & Mucchielli, 2021, p.15).

L'entretien semi-directif permet de guider partiellement la conversation des personnes interrogées en se concentrant sur différents thèmes définis en amont par l'enquêteur·ice dans un guide d'entretien : « *une partie où le sujet s'exprime librement et une partie guidée où l'interviewer oriente les verbalisations* » (Blanchet & Gotman, 2010, p.35).

La forme semi-directive d'un entretien offre à la fois une structure claire avec des points de référence pour orienter la discussion et ne rien oublier des sujets à aborder, tout en offrant une grande flexibilité. Ceci permet à chaque individu de s'exprimer librement sur son expérience dans le domaine étudié, et permet à l'enquêteur·ice d'adapter la structure en fonction des réponses. Des relances peuvent également être utilisées pour orienter la discussion vers des points non abordés ou pour approfondir un thème particulier. Blanchet & Gotman (2010) définissent trois types de relances : la « *contradiction* », qui s'oppose au point de vue du sujet, « *la consigne ou question externe* » qui introduit un nouveau thème, et enfin la « *relance* », un acte réactif qui demande à l'interviewé·e à d'explicitier davantage sa pensée.

b) Le guide d'entretien

Nous avons fait le choix d'utiliser les entretiens de deux manières différentes. Dans cette partie, les retours des professionnels serviront d'illustration de l'évolution des habitudes de travail dans la post-synchronisation. Nous avons établi un guide d'entretien comprenant des questions ouvertes, ainsi qu'une liste de thèmes à aborder. Cette dernière a été complétée progressivement au fil des différentes rencontres préalables et des observations de séances de post-synchros. Nous nous sommes appuyés sur cette liste pour orienter nos relances. Nous avons dû adapter un peu le guide d'entretien en vue des différents métiers dans la post-synchro : mixage, direction de plateau et prise de son (cf tableau p.29).

Les questions avec un fond marron sont celles qui sont destinées uniquement aux mixeur·se·s.

Dans cette partie, les thèmes qui nous concernent dans le guide sont l'évolution dans la post-synchro, les logiciels utilisés actuellement et les améliorations dans l'intégration de la post-synchro au cours des années.

Dans la partie suivante nous analyserons les questions sur les enjeux de la post-synchro. Nous allons alors nous intéresser aux enjeux de la post-synchro et aux aspects relationnels.

Thèmes	Questions principales	Relances
Présentation de la personne	Quelle est votre expérience avec les PS ?	Étape d'intervention dans la chaîne de post prod Années d'expérience, tranche d'âge, parcours
Evolution dans la PS	Comment les PS ont changé depuis le début de votre vie professionnelle ? Votre façon de travailler/communiquer avec les autres a-t-elle changé ?	Dater les événements si possible, accent sur la personne spécifique, événements Arrivée des monteur·se·s paroles en tant que directeur·ice·s de plateau Evolution des logiciels utilisés
Les enjeux de la PS	Quels ont été les cas délicats à retravailler pour vous en PS et pourquoi ? Qu'est-ce qu'une « post synchro réussie » ? Est-ce que pour vous le raccord technique et la vraisemblance sont pareils ?	Raconter des expériences précises Quand est-ce qu'on différencie à l'écoute une PS d'une prise directe ? Définissez la vraisemblance.
Relationnel	Pire et meilleure expérience de PS ?	Comment gérer des situations délicates ?
Logiciels	Quels logiciels utilisez-vous ? Est-ce qu'ils répondent à vos besoins et est-ce que vos besoins ont changé au cours des années ?	Les nouveaux logiciels sur le marché : connaissance, utilisation Quel est l'apport des nouveaux logiciels dédiés aux post-synchros ?
Améliorations dans l'intégration de la PS	Qu'est-ce qu'il manque pour améliorer le travail d'intégration des post-synchros ? Quelles contraintes rencontrez-vous souvent ?	Expérience en tant qu'utilisateur·ice, satisfaction, besoins, frustration(s)
	D'autres choses à rajouter ?	

c) Présentation du corpus de professionnels

Nous avons choisi un corpus de professionnels comprenant 5 cinq mixeurs, 3 directrices de plateau, dont une monteuse paroles ainsi qu'un perchman spécialisé dans le domaine de la post-synchro. Voici la liste de personnes interviewées ainsi que leurs identifiants¹² :

- Olivier Guillaume, mixeur (MI1)
- Matthieu Tertois, mixeur (MI2)
- Cyril Holtz, mixeur (MI3)
- Niels Barletta, mixeur (MI4)
- Aymeric Dupas, mixeur (MI5)
- Katia Boutin, monteuse paroles et directrice de plateau (MP1)
- Anne Maisonhaute, directrice de plateau (DP1)
- Catherine Taieb, directrice de plateau (DP2)
- Pierre Picq, perchman (P1)

La prise de contact avec ces personnes a été effectuée de diverses manières :

- Dans le cadre de l'école (interventions des professionnels ou parrainage)
- Dans le cadre de stages effectués chez eux ou dans les établissements qu'ils fréquentent
- Grâce aux recommandations données par d'autres professionnels

L'absence de mixeuse dans ce corpus s'explique par la facilité d'accès aux mixeurs mentionnés et le fait que ce métier est exercé par peu de femmes. Avec plus de temps, leur inclusion dans le corpus aurait été souhaitable. Nous constatons également l'absence de directeur de plateau, qui est également un résultat de facilité de contacts et d'accès. Ceci est par ailleurs la réalité dans ces métiers, et nous pensons que le résultat de l'analyse en sera peu affectée.

Nous avons choisi de mener les entretiens en présentiel dans le lieu de travail de l'interviewé-e dans la mesure du possible. Nous voulions rester en proximité d'un lieu qui leur

¹² Ces identifiants sont des références qui seront utilisés pour citer les personnes correspondantes plus loin dans le texte.

était à la fois familier et leur permettait de se souvenir de leurs habitudes de travail. Les fermetures des auditoriums de mixage étant gérées par des recorders la plupart du temps, nous n'avons pas pu rester à l'intérieur pour réaliser les entretiens car cela aurait impliqué des horaires supplémentaires pour ces derniers. Il n'aurait également pas été possible de faire des entretiens en cours de mixage à cause des contraintes de temps. Nous avons alors utilisé un espace commun non loin de l'auditorium, qui nous convenait même si l'idéal aurait été d'avoir les outils sous la main. Malheureusement, un entretien en présentiel n'était pas toujours faisable, et nous avons dû recourir à des entretiens en visioconférence pour Katia Boutin, Catherine Taieb et Niels Barletta en raison de la distance géographique et un emploi du temps chargé.

Avant de démarrer l'entretien nous nous sommes présenté.es, ainsi que le sujet de ce mémoire et cadre de l'entretien. Nous avons également fait signer un formulaire de consentement (cf. annexe p.98) qui rappelle le projet et l'enregistrement de l'entretien et donne le choix aux interviewé.es d'être anonymisé.es.

Nous rappelons que ces entretiens feront l'objet d'une analyse approfondie s'appuyant la méthode de la théorie ancrée dans la troisième partie de ce mémoire. En ce qui concerne la suite de cette deuxième partie, nous restons dans une constatation de l'état de l'art et de l'évolution des habitudes de travail qui pourra s'appuyer sur des citations des entretiens pour étayer la description de nos pratiques.

II- Du studio d'enregistrement à l'auditorium de mix

La prise de son de post-synchro a connu des évolutions techniques considérables ces vingt dernières années, qui ont changé à leur tour la méthodologie de travail, allant de la préparation de la séance jusqu'à l'étape d'intégration dans le mixage.

a) Le système de travail

- Le poste de la direction de plateau

Nous avons défini le rôle de le-la directeur·ice de plateau dans la partie précédente (cf. p.25). La plupart du temps c'est un poste assuré par une personne à part qui est introduite à l'étape de la détermination. Il y a plusieurs avantages à cela : c'est une personne avec des compétences spécifique en ce domaine qui sait répondre à la fois aux demandes techniques et artistiques. Iel comprend la difficulté de l'exercice pour les comédien·ne·s et par son expérience est capable d'inspirer et assurer des prises réussies. Ayant une oreille neuve, iel est également bien placé·e en séance de détermination pour noter les problèmes de compréhension ou des incohérences de jeu.

Un cas rare mais présent aujourd'hui en France est la prise en charge de ce poste par le-la monteur·se des directs du film. Iel est présent·e dans tous les cas en séance de détermination pour indiquer la matière indispensable à enregistrer en post-synchro, y compris les respirations et présences. L'avantage dans ce cas est d'avoir quelqu'un qui connaît parfaitement les directs et qui aura déjà effectué un travail de nettoyage et de recherche de doubles. Cela évite également de rajouter une personne supplémentaire à la chaîne de travail.

« ...je connais tellement les directs effectivement et j'ai appris tellement à écouter le jeu que je sais ce qui est juste ou pas [...] » Katia Boutin, monteuse paroles et directrice de plateau (MP1).

Katia est un exemple de monteuse des directs qui assure souvent la post-synchro dans les films sur lesquels elle travaille. Elle a étudié le système ADR en travaillant sur des films en langue anglaise et a été étonnée de découvrir que le travail de montage parole et de la supervision de post-synchro était séparé en France.

« Au début ça m'a étonnée [...] c'est quelqu'un qui ne connaît pas le film qui arrive en détermination de post-synchro et qui découvre généralement le film à cette occasion... Il me semblait un peu incongru d'avoir une personne extérieure mais ça marche très bien, c'est pas du tout une critique. »

Il y a donc des avantages aux deux systèmes et ce sont deux manières de travailler différentes qui peuvent être adaptées à certains projets. Il y a d'autres monteur·se·s des directs qui sont également appelé·e·s à superviser les post-synchros qu'ils monteront, mais cela reste rare dû aux compétences nécessaires, souvent cumulées au cours des expériences en studio.

• L'ADR en France

Katia a également catalysé l'arrivée du système ADR en France. En accueillant des productions et comédien·ne·s américaines et anglophones, ce système s'est introduit dans la post-production française. La bande rythmo reste la technique utilisée par défaut mais l'ADR offre une manière de travailler plus intuitive pour les comédien·ne·s qui aiment refaire les répliques en écoutant plutôt qu'en lisant une bande rythmo.

L'ADR n'est pas populaire aussi à cause d'une technique similaire à elle qui existait déjà avant son arrivée dans le cinéma français : la technique du « perroquet ». Elle consiste à faire écouter au comédien ou à la comédienne la réplique à refaire via un casque, puis à la répéter immédiatement après. Cela facilite la reproduction de l'intonation et du timbre de la voix.

Le côté intuitif est très apprécié pour des textes plus courts, mais la bande rythmo reste plus précise pour des phrases longues, ainsi que pour les présences et les respirations. De plus, écouter une réplique peut influencer le jeu du comédien, ce qui n'est pas souhaitable dans le cas d'une post-synchro qui vise un changement de jeu. Il n'est également pas possible de noter des respirations et des présences telles qu'avec la bande rythmo.

« Avec l'ADR tu as la musique dans la tête [...] c'est une technique qui fonctionne assez bien quand on veut refaire quelque chose de court qui doit être un peu semblable à ce qu'on avait dans le direct. » Niels Barletta, mixeur (MI4).

« Ça dépend des comédiens, il y a des comédiens qui sont très à l'aise dans un cas, pas du tout dans l'autre et à l'inverse...ça dépend s'ils ont une bonne oreille ou pas ». Pierre Picq (P1)

Dans le cas idéal, il faut pouvoir proposer aux comédien·ne·s la technique avec laquelle ils sont le plus à l'aise ; le perroquet ou l'ADR au cas où ils ont une oreille musicale et une préférence pour cette technique, et la bande rythmo si la précision de cette méthode leur plaît.

- Le travail préparatoire

Les post-synchros ont historiquement été considérées comme une solution pour corriger des problèmes techniques qui sont survenus lors de l'enregistrement des dialogues sur le tournage. Même si cela reste en partie vrai, en France la plupart des répliques post-synchronisées aujourd'hui sont dues à des problèmes de compréhension et d'articulation ainsi que des changements d'intention de jeu d'acteur.

L'évolution technique de la prise de son sur le tournage et le grand souci du cinéma français à produire le meilleur son direct possible, couplés avec les outils de nettoyage et de restauration sonore puissants donnent lieu à des répliques souvent « sauvables » même en cas de soucis techniques. La suite RX développé par iZotope est un exemple de ces outils, les différents modules permettent de retirer des bruits parasites, des bruits de fond pour avoir un meilleur rapport signal sur bruit ou encore à isoler la voix du bruit de fond. Cette modification a considérablement réduit le temps nécessaire pour enregistrer les post-synchros, tout en ajoutant une dimension artistique plus importante à cette tâche.

« Effectivement il y a des choses qu'on n'aurait pas acceptée il y a quelques années qu'on arrive à accepter là en se disant qu'on va pouvoir rattraper et garder la prise... » Katia Boutin, (MP1)

« Avant on mettait dix jours pour un film. Aujourd'hui avec le montage son, le nettoyage etc., on n'a pas besoin de faire la moitié parce que c'est sauvable. » Pierre Picq, perchman, (P1)

b) Evolution de la prise de son en studio de post-synchro

- Le passage en numérique et la libération ergonomique

La bande magnétique utilisée autrefois pour enregistrer les post-synchros permettait trois prises sur une bande. Il fallait donc changer de bande dès qu'il y avait plus de trois prises, une tâche laborieuse qui introduisait beaucoup d'inertie. La rythmo était calligraphiée à la main sur des bandes de 35 mm qui étaient projetées avec un projecteur. Le support image, la rythmo et l'enregistrement étaient sur trois supports physiques différents, et le transport¹³ était limité à 4 fois la vitesse du film ; la gestion de la temporalité était alors beaucoup plus lente.

« A chaque fois qu'on enregistrait plus de trois fois la personne il fallait changer de bande magnétique etc., c'était beaucoup plus laborieux. » Pierre Picq (P1)

« La principale (différence) c'était l'inertie du système. » Olivier Guillaume, mixeur cinéma (MI1).

« Aujourd'hui c'est ergonomiquement beaucoup plus fluide » Aymeric Dupas, mixeur cinéma (MI5).

Au début des années 2010, le système de doublage et de post-synchronisation a commencé à évoluer grâce à l'apparition de logiciels numériques capables de lire une bande rythmo et de la synchroniser avec la vidéo et l'enregistrement audio (souvent sur ProTools) grâce au Timecode. Le logiciel le plus utilisé aujourd'hui pour ces tâches est Mosaic, développé par Noblurway.

Le gain de temps grâce aux évolutions techniques a impliqué un travail plus minutieux quant aux choix des éléments à post-synchroniser. Avant l'arrivée du numérique, il était rare de faire des respirations et des présences qui ne relevaient pas du dialogue. Ces éléments étaient post-synchronisés quand ils étaient écrits dans le scénario, ou quand ils avaient une importance particulière. Aujourd'hui, il y a un rajout systématique de ces présences sur des gros plans au

¹³ Le transport est l'action de se déplacer sur la timeline d'un support audiovisuel, qui comprend des fonctions telles que le démarrage, l'arrêt, la pause et le passage à une partie différente.

cas où les directs n'en contiennent pas suffisamment. Les présences et la « vie » ajoutées au personnage permettent une meilleure intégration et vraisemblance.

« [...] on va plus chercher la petite bête aujourd'hui que bien évidemment il y a 20 ans ou 30 ans. Par exemple c'était rare de faire des respirations et présences [...] Aujourd'hui dès qu'on a un gros plan dans un film si on l'a pas dans son direct on va systématiquement l'enregistrer » Pierre Picq (P1).

Il y a également une dimension moins protocolaire en séance d'enregistrement depuis le passage au numérique. Il est désormais possible de changer l'ordre des séquences à post-synchroniser pour les comédiens, offrant ainsi une grande flexibilité sans implication de perte de temps. En effet, il y a parfois des comédien·ne·s qui se trouvent bloqué·e·s devant une réplique, et le fait de passer à une autre séquence afin de revenir plus tard à celle qui pose problème peut alors être à la fois plus rassurant pour les comédien·ne·s, plus ergonomique et efficace finalement.

Aujourd'hui, le·la mixeur·se peut faire réécouter la prise immédiatement après l'enregistrement pour vérifier la qualité de la post-synchro, ce qui aurait introduit auparavant cette inertie dont parlent les mixeurs cités ci-haut. Ce changement permet de ne plus avoir à être tributaire de l'oreille de l'ingénieur du son et ouvre l'écoute de la post-synchro enregistrée à tout le monde présent dans l'auditorium.

« Donc on a gagné à la fois cette rapidité d'exécution et l'oreille critique des différents intervenants d'une séance d'enregistrement. » Olivier Guillaume (MI1).

Une évolution qui découle de cette ouverture d'écoute est la volonté d'entendre une intégration de la post-synchro dès la prise de son. Bien que l'équipe de réalisation, de production et les acteur·ice·s aient une riche expérience dans le domaine du cinéma, ils n'utilisent pas quotidiennement les outils de la post-production sonore. Les mixeur·se·s ont la capacité de se projeter dans le résultat final en utilisant des effets tels que la réverbération et l'égalisation, tandis que les autres parties prenantes aimeraient entendre le résultat en écoutant la prise mixée pendant la séance de post-synchronisation.

« Ce qui est vraiment demandé maintenant je trouve c'est que ce soit intégré le mieux possible dès l'étape de la prise. Certains oublient qu'il y a l'étape de mix après et que si ça

pas parfaitement bien, c'est que il manque cette étape. » Matthieu Tertois, mixeur cinéma (MI2).

Pour ce faire, les mixeur·se·s préparent des réverbérations provisoires en amont pour qu'à la réécoute, la post-synchro soit acoustiquement proche du son direct. Il y a souvent une réverbération courte pour les intérieurs et une plus adaptée aux extérieurs, ainsi que des réverbérations pour des lieux très spécifiques si besoin (ex. une église, très réverbérante).

« J'ai toujours 2 réverbes, j'en choisis des suffisamment neutres [...] Une un peu longue et une plus courte pour décoller un peu pour qu'elles passent un peu partout. Et puis j'ai quelques snapshots¹⁴ que je prépare en début de post-synchro si je sais qu'il y a des lieux vraiment spécifiques... » Matthieu Tertois (MI2).

- Une meilleure intégration dès la prise de son

Avant la fin des années 90, la majeure partie du cinéma français utilisait la perche Fisher pour la prise de son de post-synchro (voir fig. 2.2.1). Cette perche est montée sur une structure et manipulée via des manivelles et d'autres mécaniques par le·la perchman ou perchwoman. La longueur du bras de la perche, ainsi que l'orientation du micro sont maniées par iel. Même si c'était une amélioration par rapport au micro sur pied, la Fisher était plus statique qu'un perchman ou une perchwoman. Cette contrainte de mobilité introduisait un jeu « figé » chez le·la comédien·ne, qui devait s'abstenir de trop bouger par soucis technique. Ceci impliquait un travail d'intégration de la post-synchro à l'étape du mixage bien plus laborieux car le mouvement devait être rajouté à cette prise de son.

« (Pendant) très longtemps (on prenait le son) avec le U87 et après avec beaucoup de travail de mixage et de correction derrière, tu pouvais faire passer ça comme du direct [...] puis y'a eu le Fisher mais tu peux plus expérimenter avec une perche. » Pierre Picq (P1)

¹⁴ Les snapshots sont des fonctions qui contiennent des paramètres que l'on peut enregistrer et rappeler rapidement, évitant ainsi d'avoir à les reconfigurer à chaque fois.



Fig 2.2.1 Une perche Fisher, jlfisher.com



Fig 2.2.2 Microphone U87, thomann.de

Depuis le début des années 2000, la prise de son de post-synchro en France se fait comme en tournage : avec une perche et un HF (micro-cravate). Souvent les mêmes modèles de micro (voire exactement les mêmes micros que ceux du tournage) sont loués à nouveau pour la séance d'enregistrement afin d'avoir une couleur et timbre de la voix la plus proche possible du son direct. Le·la perchman ou perchwoman, en accord avec le·la mixeur·se, fera ensuite de son mieux pour se placer ainsi que d'introduire des mouvements de perche si nécessaire pour que le timbre de la voix ressemble à ce qu'il entend dans le direct. Ceci incite un son moins « statique », avec une prise de son qui peut s'adapter à la valeur des plans images.

« *J'ai mon casque, je balade ma perche mon micro, je sais ce que j'entends par rapport à ce que j'ai entendu du son direct [...] Avec l'expérience j'arrive à avoir une idée de ce qui a été fait au direct...* » Pierre Picq (P1)

La volonté de recréer les conditions du tournage n'est pas limitée à la technique de la prise de son. Les comédien·ne·s sont souvent invité·e·s à se mettre en mouvement pour trouver une justesse de jeu qui raccorde avec le direct. Dans le cas d'une post-synchro lors de la marche, de la course, d'un combat etc., il est difficile d'intégrer la réplique post-synchronisée au mixage s'il manque le mouvement à la source. Ce cas est très représentatif d'un des enjeux majeurs de la post-synchro : refaire la prise de son proprement mais si elle est trop propre, sans accidents ni mouvement, elle est statique et difficile à intégrer. Il faut donc savoir doser le niveau de mouvement sans risquer une nuisance sonore.

« *Pour ce qui est de l'effort, souvent je fais faire des pompes ou des courses aux comédiens avant de faire ce genre de choses pour qu'ils soient vraiment essoufflés [...] Il faut être dans l'effort en fait sinon ça ne marchera jamais* » Katia Boutin, (MP1)

Un changement supplémentaire qui favorise la liberté de mouvement des acteur·ice·s est la suppression d'un objet qui avait une double fonction, à la fois comme repère pour le placement du micro et comme point d'appui sur lequel les acteur·ice·s pouvaient se reposer. Il s'agit de la barre de post-synchro.



Fig 2.2.3, Un comédien à la barre avec un micro fixe, La depeche, <https://www.ladepeche.fr/article/2018/01/27/2729964-acteur-jeremie-bedrune-fait-doublage-maigret-coco.html>

C'est un élément présent depuis bien longtemps en auditorium d'enregistrement ; l'acte de réaliser des post-synchronisations pour les comédiens est couramment appelé "passer à la barre" dans le milieu du cinéma. Grâce au passage à une prise de son à la perche qui suit les comédien·ne·s, iels n'ont plus besoin d'un point de repère ou d'appui et donc de limiter leurs mouvements.

« *On essaie d'enlever la barre pour justement pas que les comédiens soient trop statiques, pour qu'ils puissent bouger, être libres de leur mouvement ; c'est ultra important pour une bonne post-synchro. On a un peu libéré le côté un peu trop statique un peu trop figé de de la post-synchro.* » Catherine Taïeb, directrice de plateau (DP2).

Cependant la barre n'a pas été entièrement enlevée pour des raisons d'habitude. Elle facilite également une répartition d'énergie corporelle qui peut aider les acteur·ice·s.

c) Le montage et mixage

Il existe une multitude de logiciels qui servent à intégrer une post-synchro aux étapes de montage son et de mixage. L'étape de montage des post-synchros est essentielle au synchronisme et à la crédibilité du dialogue, et peut être chronophage du fait de la précision (à l'image près) du travail demandé.

- Le recalage de post-synchro

Une fois les post-synchros enregistrées, le·la directeur·ice de plateau ou le·la monteuse parole récupère la session ProTools pour « recalcr » les prises. Le recalage consiste à remettre en phase la perche et le HF si cela n'a pas été fait lors de la session d'enregistrement (le logiciel Auto-Align¹⁵ permettant de faire cette étape automatiquement) et surtout d'assurer le synchronisme labial des post-synchros. Les acteur·ice·s ne sont pas toujours synchrones à l'image près, et c'est une étape très importante pour l'intégration de la post-synchro et pour que cela passe inaperçu. Pour ce faire, iel va découper les prises en plusieurs parties, voire plusieurs syllabes, et les décaler, ou appliquer des outils de *stretch*¹⁶ dessus. Il est crucial que

¹⁵ Auto-Align est un logiciel de Sound Radix. Ce logiciel permet l'alignement automatique de la phase entre plusieurs sources.

¹⁶ Le stretch consiste à ralentir ou à accélérer des sons sans changer la hauteur de la voix. En recalage de post-synchro, cela permet d'atteindre une meilleure synchronisation labiale.

la synchronisation temporelle et les caractéristiques sonores de la post-synchro soient cohérentes avec celles de la prise de son originale pour que le·la spectateur·ice soit convaincu·e de l'authenticité de la scène (Avarese, 2017, p.70).

La délicatesse est de synchroniser les prises sans dénaturer le son en le découpant trop ou en appliquant trop d'outils dessus.

« Il faut vraiment qu'à l'oreille la phrase soit fluide [...] il faut arrêter quand on voit à l'image que ça passe, c'est important de respecter le jeu de la post-synchro aussi » Catherine Taieb, (DP2).

Il existe aujourd'hui des logiciels de recalage automatique comme Wordfit¹⁷, VocAlign et Revoice Pro¹⁸. Ils détectent la forme d'onde du direct, redécoupent et appliquent du *stretch* à la post-synchro afin d'obtenir un synchronisme parfait. Cependant, si la forme d'onde de la post-synchronisation diffère trop de celle de la prise directe, des artefacts peuvent apparaître. *« Un bon monteur paroles ce n'est pas forcément un bon monteur de post-synchro parce qu'il y a beaucoup de monteurs paroles qui veulent que les synthros calent exactement pile-poil soient exactement comme des directs en termes de forme d'onde [...] faut pas les traiter de la même façon. »* Catherine (DP2)

- Le mixage

Tout d'abord, nous allons nous pencher sur les outils couramment utilisés dans le mixage cinéma. Bien qu'ils puissent être utilisés pour intégrer la post-synchronisation, ce n'est pas leur utilisation la plus fréquente.

L'EQ

Dans le cadre de la post-synchro, l'égalisation (l'EQ), qui permet d'intervenir sur la bande de fréquence d'un son et d'atténuer ou amplifier certaines fréquences d'un signal audio, est un outil nécessaire pour traiter les éventuelles différences de micros entre le tournage et le

¹⁷ Wordfit est un logiciel de DAR.

¹⁸ VocAlign et Revoice Pro sont des logiciels de Synchro Arts

studio d'enregistrement ainsi que le timbre de la voix. Elle permet de raccorder ce qu'on appelle la « couleur » d'un son.

L'EQ permet d'ajouter du mouvement en plus de faire des corrections. Prenons l'exemple d'un·e comédien·ne qui tourne la tête pendant une scène. Il peut parfois être difficile pour le·la comédien·ne et le perchman ou la perchwoman de reproduire ce mouvement de manière précise en séance de post-synchro. L'EQ permet alors de filtrer la voix afin de simuler le mouvement de manière fluide. En revanche, trouver les fréquences où intervenir et ensuite recréer du mouvement de manière efficace peut prendre beaucoup de temps. Le modèle Pro Q3 de chez FabFilter est un des EQ les plus utilisés aujourd'hui.

La Compression

C'est un outil classique utilisé pour le mixage voix en cinéma, il permet de contenir la dynamique et les transitoires de la voix. La prise de son sur un plateau de tournage se fait souvent avec un Windjammer et une Rycote¹⁹ à l'extérieur, afin d'éviter une gêne à cause du vent. Dans un auditorium de post-synchro il n'y a pas de vent, et donc pas besoin de cet accessoire, la prise de son récupérant alors une partie des transitoires et fréquences aiguës qui sont légèrement atténuées par le windjammer. Il est possible d'utiliser de la compression multi bande²⁰ ou des outils comme Soothe2²¹ afin de filtrer ces fréquences (d'après discussion avec Cyril Holtz, mixeur).

La Réverbération

Nous avons vu qu'il est courant dès l'enregistrement d'utiliser une réverbération provisoire pour positionner la voix post-synchronisée dans l'espace sonore des autres éléments

¹⁹ Le Windjammer (fourrure artificielle) a été conçu pour envelopper totalement la bonnette anti-vent d'un micro et offrir ainsi une protection maximale en cas de grand vent, lors de prises de son en extérieur.

²⁰ Un compresseur multibande cible une plage de fréquences spécifique avec des bascules de bandes de fréquences individuelles qui peuvent être réglées individuellement.

²¹ Soothe2 est un supprimeur de résonance dynamique. Il identifie les résonances problématiques en temps réel et applique automatiquement une réduction correspondante.

afin d'évaluer du raccord. Cette réverbération provisoire est rarement la plus adaptée. Pour améliorer l'ergonomie du travail dans une séance d'enregistrement où le temps manque pour appliquer différents effets avant de faire écouter une prise, il est fréquent que l'ingénieur du son ait prévu deux ou trois reverbes sur la session : souvent une pour l'intérieur et une pour l'extérieur. Il n'est pas question de trouver la bonne acoustique pour chaque lieu dans le film à ce stade, c'est à l'étape du mixage qu'il faudra essayer de trouver une acoustique raccord, avec des réverbérations.

Il existe plusieurs plug-ins de réverbération (à convolution ainsi qu'algorithmique) : Altiverb développé par Audio Ease, Stratus développé par Exponential Audio, etc.

Le Pitch

Ce paramètre est lié à l'intonation de la voix. Lorsque deux prises sont combinées (cela peut être une combinaison de post-synchro et de son direct), nous constatons souvent que l'une est plus haute ou plus basse que sa voisine. Le résultat est une phrase peu convaincante avec deux parties qui ne raccordent pas, mais qui peut être corrigée avec un changement de hauteur, ou de *pitch*.

Il est important d'être attentif lorsque l'on modifie la hauteur d'une partie de la réplique, car la plupart des outils de traitement modifient également les formants, qui sont des pics dans les spectres de voyelles générés par les articulations de la bouche et de la gorge. Ces formants sont essentiels pour reconnaître les sons des voyelles et donc pour la compréhension du langage. Bien que la hauteur de la voix puisse être modifiée de manière significative, les formants doivent rester cohérents pour que le résultat sonne naturel, sinon cela peut rendre la voix artificielle et peu crédible (Purcell, 2012, p.280).

Nous venons de citer (de manière non-exhaustive) les outils plus « classiques » de mixage qui ne sont pas uniques aux post-synchros. À ceci s'ajoute des logiciels plus spécifiquement développés et utilisés pour l'intégration des post-synchros.

III- Evolution des logiciels récents

Dans cette partie nous allons nous concentrer sur les logiciels spécifiquement développés pour améliorer l'intégration des post-synchros, ou qui sont principalement utilisés à cette fin. Nous essaierons de comprendre le contexte dans lequel ils ont été développés et les besoins auxquelles ils répondent, ainsi que leurs limites actuelles. Nous nous demandons également si le gain obtenu grâce à l'évolution de ces logiciels impacte l'ergonomie ou la qualité sonore. Pour cela il est important de comprendre le fonctionnement et les propositions de ces logiciels.

Nous avons mené des entretiens avec les équipes de développement de chaque produit dans le but d'approfondir la recherche sur ces logiciels ; ils sont présentés en annexe. La présentation de ces logiciels pourra également inclure l'expérience des mixeurs·se·s interviewé·e·s.

a) Dialogue Match, Izotope : Un outil global

Dialogue Match est un logiciel développé avec le but de raccorder des bouts de dialogues entre eux ; il a une donc utilisation privilégiée dans le montage des directs ainsi que les intégrations de post-synchro. Il permet de calculer plusieurs caractéristiques sonores d'un son de référence et d'appliquer des traitements à l'audio cible afin de raccorder les deux sons. Il contient 3 modules sur lesquels l'utilisateur·ice peut intervenir.



Fig 2.3.1 Dialogue Match, capture d'écran

Le module EQ très similaire à l'*EQ Match* présent sur la suite RX d'iZotope ; il calcule l'EQ à appliquer sur l'audio cible afin qu'il se rapproche de la couleur et du timbre de l'audio de référence. Il n'est pas configuré pour être extrêmement précis. L'objectif est plutôt d'ajuster le ton global du signal plutôt que de rechercher une correspondance exacte entre deux prises, car généralement, l'utilisateur·ice travaille sur deux prises (référence et cibles) très courtes, mais il va les appliquer à de nombreuses autres prises de voix, qui ont donc des caractéristiques légèrement différentes.

Le module d'ambiance match est également similaire à celui présent dans RX. Dans le cas de Dialogue Match, il extrait l'ambiance de fond présente dans l'audio de référence et l'ajoute à l'audio cible. Pour l'instant, il est seulement possible d'extraire des ambiances statiques, sans possibilité de recréer des ambiances avec des sons périodiques (par ex. des oiseaux, des gouttes d'eau, etc.). Les développeurs aimeraient ajouter cette fonctionnalité à une version future de Dialogue Match.

Le module de *Reverb Match* permet « d'extraire » la réverbération de l'audio de référence et de l'appliquer à l'audio cible. Cette fonction a été d'abord développée par Exponential Audio²², qui a été achetée par iZotope. Il n'y a pas de création d'IR²³ (réponse impulsionnelle) lorsqu'un audio référence est renseigné dans Dialogue Match. Il s'agit plutôt d'un algorithme qui cherche la réverbération la plus proche d'une base de données de réverbérations *room*, *plates* et *hall* dans le logiciel. Ce n'est donc pas la réverbération exacte qui est extraite mais une réverbération « sœur » la plus proche possible. Il est souvent nécessaire de retoucher à cette nouvelle réverbération afin d'avoir un résultat convaincant.

Dialogue Match utilise l'apprentissage automatique pour améliorer les estimations de la réverbération. Son algorithme écoute un dialogue avec de la réverbération et sélectionne un pré-réglage de réverbération très similaire. Pour ce faire, de nombreux dialogues sans réverbération ont été pré-traités avec de la réverbération artificielle par les développeurs. Ensuite, les chercheurs en apprentissage automatique ajustent l'architecture à laquelle ces enregistrements sont soumis. L'algorithme cherche à trouver le pré-réglage le plus proche et reçoit une note évaluant sa proximité par rapport au résultat souhaité. Il s'efforce ensuite d'améliorer cette note par la répétition et l'ajustement de ses paramètres.

La *reverb-match* est le module le plus questionné par des professionnels et recherché par iZotope pour deux raisons majeures. La première est sa précision, ce qui est lié à une difficulté de compréhension des *early reflections*²⁴ ou premières réflexions par l'IA (l'intelligence artificielle). « *Les premières réflexions, en particulier celles de très courte durée, sont perçues comme une EQ. Nous avons donc essayé de couvrir une grande partie de cela dans notre fonction d'EQ match. Cependant, il y a ces réflexions légèrement plus tardives qui semblent être très délicates à capturer.* » Nick LaPenn, chef de l'équipe de développement de Dialogue Match. La difficulté s'expliquerait en partie par le fait que la différence entre une

²² Développeur de plug-in de réverbération acheté par iZotope en 2019.

²³ En traitement du signal, la réponse impulsionnelle d'un processus est le signal de sortie qui est obtenu lorsque l'entrée reçoit une impulsion, c'est-à-dire une variation soudaine et brève du signal. La réponse impulsionnelle se présente sous forme d'un fichier audio contenant une captation (prise virtuelle) des caractéristiques sonores d'un environnement acoustique ou d'un instrument.

²⁴ Les "early reflections" ou font référence aux premières réflexions sonores qui se produisent peu de temps après l'émission d'un son. Elles jouent un rôle crucial dans la perception de l'environnement sonore et contribuent à notre compréhension de l'acoustique d'une pièce ou d'un espace extérieur.

queue de réverbération de 5 secondes et une réverbération de 5,1 secondes n'est pas nécessairement très importante pour notre perception. Cependant, dans le cas des premières réflexions, même un dixième de milliseconde peut modifier la perception de l'égalisation d'un signal. Nous sommes donc beaucoup plus sensibles à l'effet des premières réflexions sur un signal.

D'après les développeurs, la meilleure manière d'assurer un résultat cohérent avec Dialogue Match est de capturer et comparer plusieurs versions des audios cibles et références. Il est toujours mieux de capturer un bout de dialogue avec l'ambiance du fond même si le module ambiance ne sera pas utilisé, afin de donner une idée plus précise au logiciel pour l'EQ. Plus nous donnons des éléments à apprendre, meilleur sera le résultat ; il faut s'assurer que ce qui est capturé représente l'EQ, réverbération et ambiance moyenne de la référence et de la source.

Les temps accordés à la post-production sonore d'un film ayant été raccourcis au cours des dernières décennies, les outils permettant une bonne ergonomie tout en assurant la qualité sonore sont d'autant plus appréciés. Une prochaine version de Dialogue Match pourrait contenir des contrôles pour le pitch de la voix, outil très utilisé en intégration des post-synchros. L'approche d'iZotope suscite des interrogations parmi de nombreux professionnels, d'après nos entretiens. Certains préférant l'utilisation de logiciels spécialisés pour chaque fonction afin de garantir une qualité sonore optimale, tandis que d'autres optent pour des logiciels "tout-en-un" qui répondent à leur besoin de rapidité d'exécution.

b) Chameleon, Accentize : le *reverb-matching*

Chameleon, développé par Accentize est un logiciel pour le « reverb-matching ». C'est un outil permettant de modéliser la réverbération de tout enregistrement source. Accentize a été le premier à introduire un outil de ce genre aussi bien reçu dans le monde de la post-production.

Le principe de fonctionnement de Chameleon est simple : il suffit de fournir six secondes d'audio source. À la suite d'une analyse, le logiciel nous présente la réponse impulsionnelle de l'espace où se situe le son source. Cette réponse peut ensuite être appliquée sur un son différent en tant que réverbération à convolution.

Chameleon a été développé en utilisant le « deep learning »²⁵. Il y a un réseau de neurones (neural-network) qui est formé sur de nombreux exemples d'enregistrements, surtout du dialogue dans différentes pièces. Cet enregistrement est saisi en entrée et ensuite, le résultat souhaité (la réponse impulsionnelle ou IR recherchée) est montré au réseau neuronal. Le réseau doit apprendre le mappage (ou la correspondance) entre l'entrée et la sortie. Cet apprentissage se fait en boucle pour des milliers d'enregistrements sources ; il faut quelques jours pour que le réseau acquière les connaissances nécessaires à chaque mise à jour du produit. « Une fois l'étape d'apprentissage terminée, on passe aux tests. Il faut lui fournir de nouveaux enregistrements d'où il pourra extraire des réponses impulsionnelles, et les comparer aux résultats attendus. » Benjamin Graf, Accentize

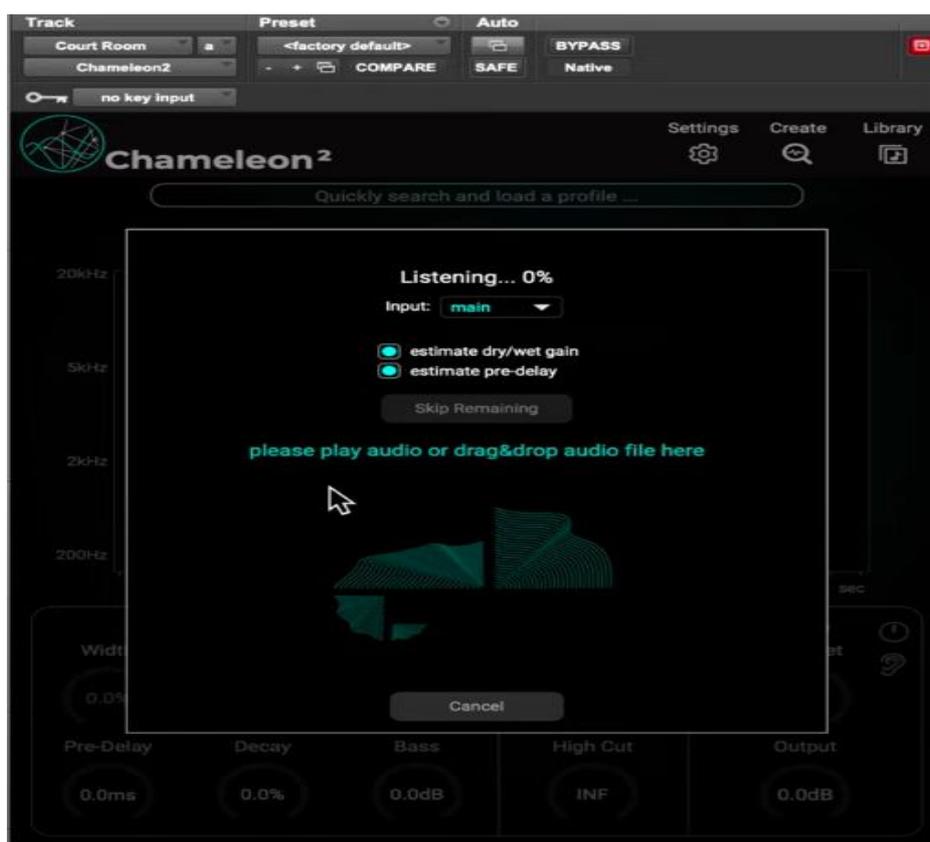


Fig 2.3.2 Chameleon2, capture d'écran de l'interface

Contrairement à Dialogue Match, Chameleon crée à chaque fois une IR différente selon l'entrée qui lui est donnée. Ceci résulte en une réverbération plus précise et proche de ce qui

²⁵ Le Deep Learning est une technique d'apprentissage automatique (Machine Learning) qui utilise des réseaux de neurones artificiels profonds (à plusieurs couches) pour apprendre à partir de données.

est recherché, et le calcul rapide des IR joue également dans l'ergonomie puissante de Chameleon.

Il est également possible d'extraire l'IR fournie par Chameleon dans un fichier audio pour l'intégrer dans un autre plug-in de réverbération (comme Altiverb) si souhaité. Rester sur Chameleon présente l'avantage d'un gain d'espace de stockage, ainsi que d'une portabilité facile des IR. Il y a un réseau neuronal dédié à la compression et décompression de l'IR, qui permet un stockage en format compressé et des rappels de réglages plus efficaces. De plus, lorsqu'une session de travail est transférée vers une autre machine, les réponses impulsionnelles ne nécessitent pas de transfert séparé. En effet, elles font déjà partie intégrante du projet et sont incluses dans les données du projet.



Fig 2.3.3 Chameleon2, capture d'écran

L'interface de Chameleon2 se présente comme ci-dessus. L'utilisateur·ice peut agir sur certains paramètres comme avec d'autres logiciels de réverbérations à convolution (ex. *Altiverb* d'*Audio Ease*). Cette réverbération est très appréciée des mixeur·se·s aujourd'hui pour intégrer les post-synchros grâce à son ergonomie et la qualité sonore qu'elle apporte.²⁶ Elle est stable dans ses résultats et son fonctionnement dans les décors d'intérieurs. Les extérieurs peuvent nécessiter quelques essais afin d'obtenir un résultat convaincant ; nous avons constaté que ces acoustiques d'extérieurs étaient les reverbes les plus dures à modéliser.

Chameleon2 permet de faire des automatisations de reverbe, chose qui est très importante dans l'ergonomie et le workflow du mixage cinéma. Chameleon3 introduira des changements algorithmiques qui ajouteront un élément supplémentaire à contrôler : les *early reflections*. Cette nouveauté pourrait notamment améliorer l'estimation des réverbérations extérieures, tout en permettant à l'utilisateur·ice d'avoir la main sur les réglages de ces premières réflexions afin de mieux raccorder la reverbe.

c) **Boom-It, HAL : L'ajout de mouvement**

Boom-It est un logiciel développé par HAL, une société de professionnel·le·s de l'audiovisuel. Il a pour but de simuler l'influence de mouvements de microphones sur leur captation sonore afin de reproduire les détimbrages et les imperfections qui sont présents en prise de son sur un tournage, mais pas forcément en studio de post-synchro (surtout dans le cas d'une prise de son sans perchman ou perchwoman, avec un microphone fixe.)

Il permet de simuler plusieurs micros de tournage différents ainsi que tous les mouvements de perche possibles en plus des mouvements de personnages par rapport au micro, ce qui a nécessité un travail de prise de mesures précis.

²⁶ D'après discussions auprès de plusieurs professionnels, ainsi que le feedback donné aux développeurs.



Fig 2.3.4, Boom-It, image prise sur le site de HAL-audio

La figure ci-dessus illustre les différents modules proposés par Boom-It.

Le *studio mic* et le *set mic* sont conçus pour éliminer la coloration du microphone utilisé pour l'enregistrement sonore (au choix parmi plusieurs options de micros de studio disponibles, mais pas de micros de tournage), puis ajouter la coloration d'un autre microphone (cette fois-ci de tournage). Les micros proposés dans chaque cas sont soigneusement choisis, car ce module Boom-It est principalement destiné à être utilisé dans le cas où la prise de son de post-synchronisation est effectuée avec un micro fixe (comme le Neumann U87 sur pied), afin d'ajuster la couleur sonore à celle du micro utilisé lors du tournage. Cependant, il serait risqué de permettre aux utilisateur-ice-s d'utiliser n'importe quel microphone en entrée et en sortie dans ce module, car il est plus facile de détériorer un son que de l'améliorer. Il serait par exemple difficile de transformer un son enregistré avec un microphone canon en un son enregistré avec un U87.

Il est aussi possible d'émuler le son et les mouvements de tête par rapport aux micros-cravates. Ce module permet également l'ajout d'une bonnette anti-vent virtuelle au micro, accessoire souvent présent en tournage, qui adoucit légèrement le son.

Le module *head tilt and angle* permet de simuler la couleur et le timbre de la voix à différentes positions relatives par rapport au microphone. Pour créer ce module, l'équipe de développement a utilisé un simulateur de torse équipé d'un haut-parleur à la place de la bouche, émettant des impulsions sonores enregistrées à différentes distances et angles à l'aide d'une perche.



Fig 2.3.5, Simulateur de torse utilisé pour les mesures de Boom-It
Photo prise sur le twitter de Hal Audio

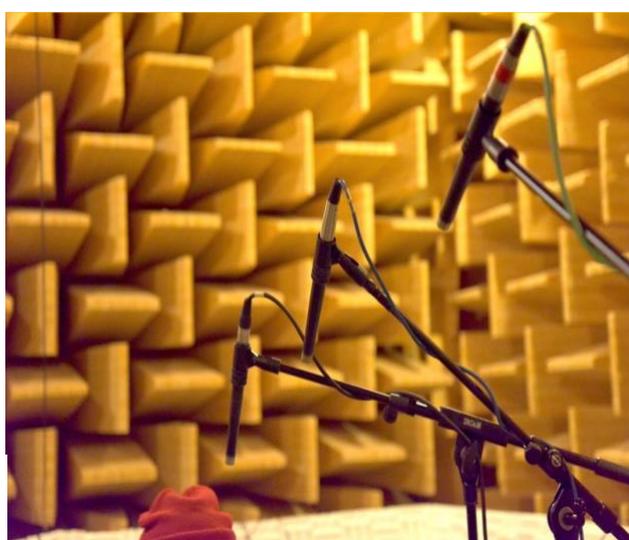


Fig 2.3.6, placement des micros
Photo prise sur le twitter de Hal Audio

Pour le *head angle*, des mesures ont été effectuées à des intervalles de 10° dans l'axe horizontal, couvrant un angle de 180° . Pour le *head tilt*, des mesures ont été prises avec les positions de la tête inclinées à -30° , 0° et 50° dans l'axe vertical. « Les réponses impulsionnelles ont ensuite été converties en courbes d'EQ (qui ne sont pas à phase linéaire), et pour les

positions intermédiaires à celles mentionnées le logiciel utilise des modules d'analyse afin de passer d'un EQ à l'autre de manière fluide. » Antoine Martin, équipe de développement chez Hal-Audio. Le module distance a été également réalisé grâce à des mesures similaires prises à des distances allant de 30cm à 5m.

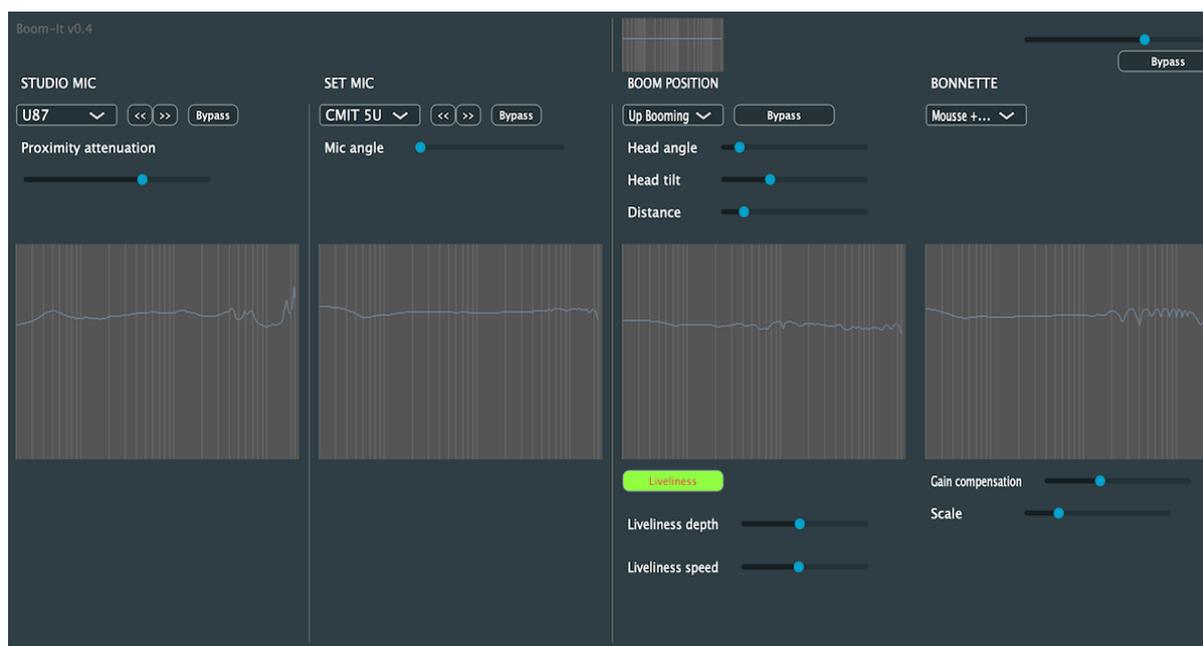


Fig 2.3.7, Version de travail Boom-It, HAL

Nous pouvons observer les EQ complexes que Boom-It calcule et applique selon les positions de la tête par rapport au micro. À la sortie du logiciel, nous aurons la sommation des EQ appliqués par chaque module dans l'ordre présenté dans la figure ci-dessus.

Le *liveliness* : il s'agit d'un paramètre proposé par Boom-It qui est assez nouveau et très intéressant pour notre sujet. Le *liveliness* permet d'ajouter des micro-mouvements qui sont induites de manière aléatoire dans le *head tilt*, le *head angle*, ainsi que la distance, et qui simulent une prise de son de tournage où le perchman et l'interprète ne sont jamais parfaitement statiques.

« C'est un logiciel qui produit de l'aléatoire et de la complexité. C'est recherché depuis très longtemps parce que la complexité et l'aléatoire c'est justement ce qui fait la vie » Olivier Guillaume, mixeur, sur Boom-It.

Le dernier module est plus classique et permet d'avoir la main sur le gain, le gain compensation ainsi que le « scale » qui est comparable avec le dry/wet²⁷ d'un effet.

Boom-It, étant un logiciel très récent (sorti en novembre 2022), est encore en cours de découverte par les professionnels. Il a le mérite d'avoir appliqué la recherche des imperfections qui donne cette « vie » au son direct de plateau que l'on a évoquée plusieurs fois au cours de cet écrit. Ce logiciel est l'équivalent d'un EQ et peut-être utilisé de manière similaire dans une session de mixage.

Un élément important qui a émergé des discussions avec les développeurs de logiciels ainsi qu'avec les professionnel·le·s interviewé·e·s concerne l'impact croissant de l'intelligence artificielle (IA) et de la synthèse vocale sur l'industrie audiovisuelle. Nous avons vu l'utilisation du *machine learning* dans le cas de Chameleon et Dialogue Match. Un exemple notable de la synthèse vocale est Voice Designer, développé par Soundly, qui permet de générer différentes voix à partir d'un texte (voir annexe p.108). De plus, il existe des logiciels qui permettent d'analyser et de modifier une voix pour la transformer en une autre, une technologie qui a été utilisée par Le Monde, avec l'aide de l'IRCAM, pour reconstituer l'Appel du 18 juin du général de Gaulle²⁸. Les nouvelles perspectives qu'amènent ces technologies ne font pas l'objet de ce mémoire, nous noterons néanmoins que l'on peut s'attendre à ce qu'elles aient un fort impact dans le futur, sur les pratiques, les métiers et l'économie de la post-synchronisation, et d'ailleurs de la post-production dans son ensemble.

²⁷ Lorsqu'un effet est appliqué à un signal audio, le signal d'origine non modifié est considéré comme le signal "dry", tandis que le signal modifié par l'effet est appelé le signal "wet". En ajustant le paramètre entre ces deux extrêmes, on peut obtenir différentes combinaisons de signal sec et signal modifié, permettant de doser l'intensité de l'effet dans le mixage audio.

²⁸ Video sur la chaine YouTube du journal Le Monde : <https://youtu.be/-LEg3TU9-kU>

Partie 3 : Les enjeux de la post-synchro, analyse en théorie ancrée

I- Le contexte des entretiens

« *L'entretien est un instrument d'investigation spécifique, qui aide donc à mettre en évidence des faits particuliers [...] dont la parole est le vecteur principal* » (Blanchet & Gotman, 2007, p.25).

Nous l'avons précisé dans la partie précédente : la post-synchro ne fait pas l'objet d'une littérature assez riche pour nourrir ce mémoire. Cette partie consiste à analyser des questions posées lors de l'entretien afin de mieux comprendre certains concepts, notamment la vraisemblance et la notion d'une « post-synchro réussie ».

II- Principe de l'analyse thématique

a) L'analyse thématique par la théorie ancrée

« *L'analyse thématique défait en quelque sorte la singularité du discours et découpe transversalement ce qui, d'un entretien à l'autre, se réfère au même thème.* » (Blanchet & Gotman, 2010, p.97)

En 1967, Glaser et Strauss ont inventé la méthode de l'analyse par théorisation ancrée, également appelée *grounded theory* (Glaser & Strauss, 2017). Cette méthode consiste à analyser de manière itérative et progressive un phénomène, sans préjuger de son contenu ou de l'ordre des éléments qui le composent, contrairement à l'analyse de contenu. Nous cherchons à relever des thèmes qui regroupent les propos mis en avant par nos interviewé-s, afin de les mettre en lien, suite auxquels nous établissons une hiérarchie de catégories et de sous-catégories.

b) La retranscription

Nous avons choisi d'effectuer des enregistrements audio des entretiens avec l'accord des participants afin de procéder à une retranscription à l'aide de la dictée automatique de

Word. Dans le but de garantir une lecture fluide, nous allons faciliter la compréhension en éliminant les hésitations et les formulations orales des participant·e·s. Une relecture de chaque entretien nous a permis de catégoriser, puis de sous-catégoriser les idées présentées par les participant·e·s

III- Analyse des résultats : Les enjeux de la post-synchros

a) Les catégories

Nous avons dégagé trois catégories dans les entretiens sur une bonne intégration de post-synchro. Le nombre de personnes qui nous ont parlé de chaque sous-catégorie est indiquée dans le tableau ; cela nous donne une idée de la fréquence à laquelle les participant·e·s évoquent les idées listées.

Catégories	Sous-catégories	Mentionné par	Nb
La séance d'enregistrement	La communication avec les comédiens	P1, MP1, DP1, MI1, MI2, MI4	6
	Différents cas pratiques	MP1, DP1, DP2, MI1, MI2, MI3, MI4, MI5	8
	L'émotion et le mouvement : raisons de préférence pour le son direct	MI2, MI3, MI5, DP2	4
Une post-synchro réussie	La recherche aboutie en séance d'enregistrement	P1, MI2, MI3, DP1, MI4	5
	L'aspect technique	MI1, DP1, MP1, MI4	4
	Le respect du sens du film	MI1, MI3, P1, MI5	4
La vraisemblance recherchée en post-synchro	Le lien avec le son direct	MI1, MI2, MP1, MI3, MI4, DP1, DP2	7
	L'importance du jeu d'acteur et l'aide au jeu	P1, MI1, MI2, MP1, DP2, MI3	6
	Simulation des conditions réelles : l'aléatoire et l'accident	P1, MI1, MI2, MI3, MP1, DP2	6
	Au-delà du raccord technique : la vraisemblance, une sensibilité humaine	P1, MI1, MI3, MI4, MP1, DP2	6

À partir de ces sous-catégories relevées par l'analyse nous avons fait des regroupements thématiques pour certains d'entre eux. Ainsi, la sous-catégorie « différents cas pratiques » de la catégorie « la séance d'enregistrement » abordera quatre sous-thèmes :

- Les présences et respirations
- Les changements d'intention
- Jauger la portée de la voix : l'utilisation du casque
- Les enfants

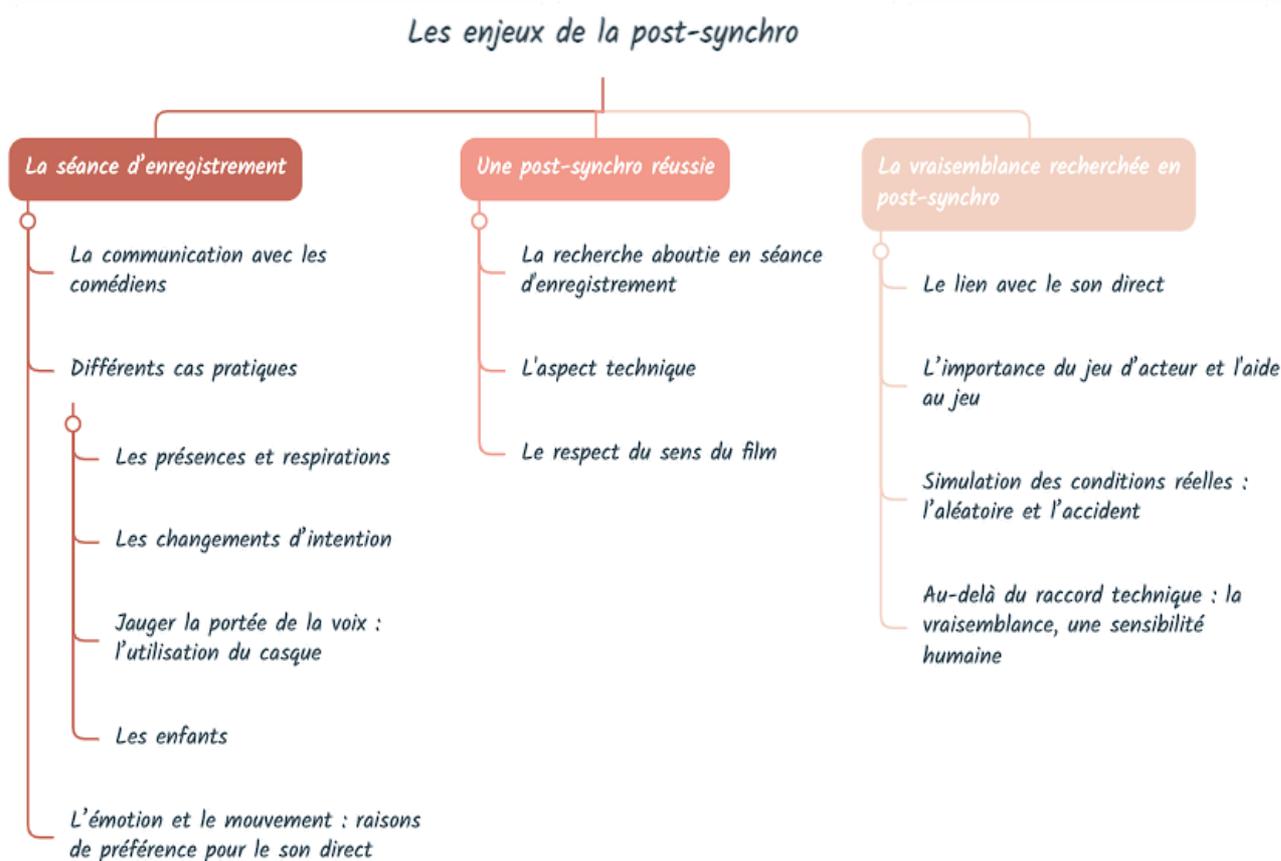


Fig 3.3.1 L'arbre thématique de l'analyse des enjeux de la post-synchro

b) La séance d'enregistrement

- La communication avec les comédiens

La gestion des aspects humains et des relations est l'un des aspects les plus complexes de la post-synchronisation. Il est crucial que le·la réalisateur·ice et l'acteur·ice se comprennent et s'entendent pour éviter les problèmes de communication. Certains acteur·ice·s peuvent avoir du mal à retrouver l'émotion du direct lors d'une post-synchronisation, mais d'autres voient cela comme une opportunité d'améliorer leur performance, car une post-synchro, c'est une prise de plus, donc une chance de plus.

En France, il n'y a pas de formation offerte aux comédien·ne·s pour s'entraîner à la post-synchro contrairement aux écoles en Angleterre ou aux Etats-Unis, où la post-synchro est vue comme la continuation du travail de tournage. La plupart des comédien·ne·s découvrent l'exercice en travaillant. Il est nécessaire de les mettre à l'aise pour qu'il puisse jouer comme sur un tournage dans le contexte particulier et peu naturel de l'auditorium de post-synchro.

Une des manières de faire cela est en prenant en compte l'ordre de déroulement des scènes à refaire, car cela peut influencer l'ambiance de toute la séance avec un·e comédien·ne. Il serait plus encourageant pour les acteur·ice·s de commencer par des ajouts de textes ou des respirations plutôt que de refaire des scènes existantes pour des raisons techniques ou de jeu. Iels peuvent avoir l'impression de refaire le film car le·la réalisateur·ice n'est pas satisfait de leur jeu. Il peut être nécessaire de commencer par des exercices plus simples pour mettre les acteur·ice·s en confiance.

- Le seul truc qui puisse mal se passer c'est que le réalisateur et l'acteur ne s'entendent pas donc ils ne se comprennent pas. (P1)
- Si c'est une posture rétissante ou quelqu'un qui ne soit pas trop en confiance avec le dispositif, il faut peut-être commencer la séance par des choses très simples à faire : des respirations, des choses qui vont mettre déjà dans le bain de ce contexte particulier du studio d'enregistrement qui n'est pas naturel pour les comédiens. (MI1)

- Il y a beaucoup de gens qui arrivaient en disant « pourquoi je refais ça moi je n'arriverai jamais à me remettre dans l'émotion du direct » et je comprends que ce soit dur mais en même temps ça fait partie de leur travail. (MP1)
- Les Anglais arrivent avec l'idée de dire « attendez je crois que je peux faire mieux » et ça c'est génial. (MP1)
- S'ils viennent et qu'ils ont l'impression de refaire le film ça peut les décourager un peu. (MI4)

Il est crucial de préserver la voix des acteur·ice·s en réservant les prises les plus exigeantes pour la fin de la séance. Lorsqu'il s'agit de scènes de forte émotion nécessitant des efforts vocaux intenses tels que des cris ou des pleurs, il est préférable de les programmer en dernier. Cela permet d'éviter la fatigue des cordes vocales et donne à l'acteur·ice l'occasion de s'adapter progressivement à l'exercice en commençant par des séquences moins complexes.

Chaque comédien·ne a sa propre façon de travailler et des facilités, il est donc essentiel de s'adapter à leurs besoins et de trouver la meilleure approche pour chacun. Il y en a par exemple qui ont une oreille musicale et une plus grande facilité avec la méthode du « perroquet » et peuvent répéter des phrases courtes qu'ils entendent au casque. Ceci est souvent le cas chez de jeunes comédien·ne·s. D'autres peuvent préférer la bande rythmo pour sa précision.

- C'est essentiel considérer leur voix comme une musique donc c'est important que l'instrument se chauffe. Il faut se garder les choses très compliquées et qui peuvent abîmer la voix pour la fin de la séance. (MI1)
- Ce qu'il faut très vite trouver c'est la meilleure façon de bosser avec tel ou tel comédien ; il y en a qui vont avoir une bonne oreille, ou d'autres qui lisent la rythmo parfaitement...(MI2)

Techniquement, il peut y avoir des phrases qui vont très vite et des comédiens qui ont du mal à articuler suffisamment, ce qui entraîne des erreurs récurrentes au même endroit. Cela peut être frustrant pour eux. Parfois, il est possible de simplifier une phrase en se concentrant sur le moment clé, plutôt que de dire une longue phrase qui n'est pas essentielle. L'important est d'avoir le bon rythme et l'énergie appropriée pour assurer la cohérence.

- Bon techniquement des phrases compliquées sont des phrases qui vont très vites (...) c'est un peu désespérant pour les comédiens. (MP1)
- On peut ne pas être rigide et simplifier une phrase (...) pourquoi dire 10 mots si peut-être on peut en dire 5 ? Ce qu'il faut raccorder c'est l'énergie et le débit. (DP1)

- Différents cas pratiques

En abordant ce sujet, nos interviewé·e·s ont tous partagé des exemples de situations délicates auxquelles iels ont été confronté·e·s. Nous avons regroupé ces cas dans différentes catégories afin de présenter les défis auxquels iels font face, ainsi que les approches utilisées par ces professionnel·le·s.

Les présences et respirations

Les présences, notamment les respirations, sont souvent considérées par les professionnel·le·s comme des éléments importants pour donner vie à la performance sonore, même s'il s'agit de respirations minimales. La gestion de ces éléments est un aspect essentiel mais délicat de la post-synchronisation. Il faut être précis pour éviter que cela paraisse artificiel (bien suivre l'image pour respecter la bouche ouverte et fermée), et ajuster le rythme naturel des respirations peut demander un travail minutieux de montage.

Un point important est la différence de perception entre les acteur·ice·s et l'ingénieur du son lorsqu'il s'agit des respirations. Il peut arriver que l'acteur·ice pense avoir effectué une respiration adéquate en termes de niveau sonore, alors que l'ingénieur du son estime qu'elle n'est pas suffisante. En effet, ce qui paraît fort au comédien peut être inaudible une fois écouté avec tous les autres éléments sonores de la scène. Il est donc essentiel de prendre en compte la différence entre un lieu de tournage et l'environnement insonorisé et silencieux, qui pousse à une tendance de retenue.

- Un cas classique (...) ce qui est le plus dur ce sont les présences, bouche ouverte, bouche fermée, pour moi les présences c'est la vie en fait et il faut être très précis là-dessus sinon ça fait faux. (MP1)
- Ça a l'air de rien mais pour arriver à coller au rythme naturel des respirations c'est un travail de dingue et de montage après pour que ça ait l'air vrai...(MI3)
- Si je demande une respiration plus forte et que le comédien a un doute, je demande au réalisateur qui est à trois mètres s'il a entendu. S'il ne l'a pas entendu à cette distance du comédien dans un auditorium insonorisé, il m'en faut plus. (MI2)

Les changements d'intention

Une des situations les plus difficiles en prise de son et en mixage de post-synchronisation survient lorsqu'il y a des changements d'intention significatifs, ce qui complique l'harmonisation du texte et des intentions avec l'image existante. Ça arrive lorsque l'on souhaite contrecarrer l'intention originale ou faire jouer des éléments qui n'ont pas été interprétés initialement. Parfois, il est nécessaire de faire passer des messages ou des émotions qui ne sont pas clairement visibles à l'écran, ce qui peut créer une distanciation entre l'image et les dialogues enregistrés.

Un exemple serait lorsque l'on demande à l'acteur·ice de timbrer sa voix en post-synchro alors qu'iel chuchotait sur le tournage, afin de faciliter l'intelligibilité dans une scène bruyante. L'intégration de nouveaux dialogues peut alors poser des défis, car l'énergie dans l'image et dans le son ne correspondent plus parfaitement.

- Les cas qui pourraient être difficiles c'est quand on a un texte qui n'est pas complètement dedans, quand on a des vrais changements de d'intention. (MI5)
- Ce genre de situation où on rajoute du texte qui n'était pas prévu à l'origine, l'intégration parfois elle peut être un peu compliquée. (MI4)

- Quand il y a trop de différences et quand on veut faire jouer à la post-synchro quelque chose qui n'est pas joué à l'image quand on veut aller contre l'intention de départ, c'est compliqué. (MI3)

Jauger la portée de la voix : l'utilisation du casque

L'un des défis consiste à s'assurer d'une bonne portée de voix chez le·la comédien·ne. Par exemple, dans le cas où iel doit chuchoter ou parler doucement, il faut quand-même maintenir une portée de voix suffisante pour un bon mixage sonore. Parfois, il est nécessaire de sortir de ce qui semble être naturel pour faciliter l'intégration de la post-synchro au mixage, car un enregistrement trop réaliste peut ne pas être une bonne idée lorsque d'autres éléments sonores sont présents.

Il est parfois également difficile de maintenir l'énergie jusqu'à la fin de la phrase, car il arrive fréquemment que l'énergie retombe. Il est important de rappeler aux acteur·ice·s de penser à leur phrase jusqu'au bout et d'éviter les décrochements, afin de garantir une intelligibilité et un bon raccord avec ce qui suit. De plus, certains acteur·ice·s ont du mal à jouer sans la présence de leur partenaire, ce qui peut être un obstacle.

Un moyen de remédier à un manque d'énergie de la part de l'acteur·ice lors de l'enregistrement est de lui fournir le son du film dans un casque. Ainsi, s'iel semble manquer de dynamisme, il est possible d'augmenter progressivement le volume du casque pour le·la stimuler et l'encourager à porter plus sa voix. Iel peut également avoir accès aux répliques des autres comédien·ne·s de cette manière. La question de la repisse²⁹ du casque se pose, mais le niveau de voix de l'acteur·ice masque généralement la majeure partie du bruit provenant du casque fermé. De plus avec les outils de nettoyage à notre disposition, il est préférable d'avoir un son légèrement parasité et un jeu juste plutôt qu'un son propre et un jeu non utilisable.

Le·la réalisateur·ice peut également s'opposer à une projection plus forte de la voix ; une astuce dans ce cas serait de faire écouter la prise précédente légèrement en dessous du reste du mix pour le convaincre (ou plus fort dans le cas contraire).

²⁹ La « repisse » fait référence au phénomène où une partie du son émis par un casque audio est capté par les microphones environnants.

- Il faut faire comprendre cette histoire de chuchoter-porter parfois pour être vraiment dans un jeu assez confidentiel intime mais qu'on ait suffisamment d'énergie pour que ça se mixe bien. (DP2)
- La difficulté c'est de leur faire terminer la phrase pour que l'énergie ne retombe pas à la fin parce que soit ça devient peu intelligible soit en mixant ça va être en dessous. (DP1)
- Parfois faire un truc parfaitement naturel, pendant des scènes de chuchotement dans un bar bruyant par exemple, alors qu'au son il va y avoir beaucoup de choses, ce n'est pas une bonne idée ; il faut parfois forcer un peu les choses. (MI2)
- Ce qui peut bloquer parfois aussi c'est qu'il y en a qui n'arrivent pas à jouer sans l'autre comédien, on peut lui mettre un casque dans ces cas. (MI3)
- Il suffit de faire réécouter ensuite la prise qui était trop faible et de la mettre trop faible en volume, généralement ils sont d'accord pour refaire. (MI2)

Les enfants

Travailler avec des enfants en post-synchronisation peut être compliqué. Les enfants très jeunes ne pouvant pas lire la bande rythmo, ils nécessitent une direction en temps réel pour ajuster leur performance. Ils ont du mal à analyser et à réagir avec leur voix de la même manière que lors de l'enregistrement en direct. Par ailleurs, la mue de la voix chez les adolescents pose un gros problème, car il est souvent difficile, voire impossible, de rattraper les changements de voix.

Effectivement, lorsqu'il s'agit de diriger de jeunes comédiens, il peut être utile de recourir à des astuces pour recréer une certaine réalité proche du tournage.

- Ce qui peut être très compliqué c'est avec les enfants qui ne peuvent pas lire la rythmo donc ça demande de les diriger en temps réel. C'est compliqué pour eux d'analyser et de faire réagir leur voix de la même manière que sur le son direct. (DP2)

- ...Un gros problème avec des ados qui muent, c'est très compliqué voire impossible à rattraper. (MI1)
- Il faut trouver des astuces aussi en dirigeant le jeune comédien par exemple en se mettant à l'autre bout de l'auditorium et en lui demandant de parler à la personne au bout. Il y a des techniques comme ça pour reproduire une espèce de pseudo-réalité plus proche du tournage. (MI1)

- L'émotion et le mouvement : raisons de préférence pour le son direct

Les séquences émotionnelles sont souvent très difficiles, car il faut se replonger dans l'émotion de la scène sans avoir son partenaire de jeu à ses côtés, dans un auditorium où l'on ne connaît pas forcément les autres personnes présentes, à part le·la réalisateur·ice. C'est une situation délicate pour toutes les personnes présentes. Les émotions extrêmes sont particulièrement délicates à interpréter, car elles nécessitent une grande énergie, et en auditorium de post-synchro il y a le défi ajouté du raccord et du synchronisme. Il vaut mieux prévoir beaucoup de temps pour ce genre de séquence.

- Il y a évidemment toutes les séquences d'émotion qui sont très difficiles parce qu'il faut se remettre dans l'émotion de la scène tournée alors qu'on est tout seul. (MI3)
- Ce qui est délicat ce sont les émotions extrêmes, c'est crier, pleurer, là il faut de l'énergie. (MI2)

Éviter un jeu statique est un enjeu majeur de la post-synchro ; nous allons en parler plus en détail dans l'analyse de la vraisemblance. Il peut être utile de faire bouger les acteur·ice·s et de les placer dans des positions correspondant à ce qui est montré à l'écran. Parfois, les encourager à se lever même si cela ne figure pas dans la scène peut contribuer à créer une sensation d'effort réel qui ne peut être transmis en restant immobile.

- Les faire bouger toujours dans la position qui correspond le mieux à (...) il y a un effort qu'on n'arrive pas à faire en étant statique. (MI2)

- Si on parle vraiment technique c'est sûr qu'un acteur qui va énormément bouger à l'image, qui va danser, ça va être compliqué parce qu'il va falloir reproduire ces mouvements. (MI3)
- En termes de raccord c'est plutôt les scènes extérieures qui sont plus difficiles et les séquences en mouvement. (DP2)

Ces deux paramètres semblent être deux raisons majeures de la préférence du son direct ; raccorder l'émotion et le mouvement est un exercice difficile à la fois psychologiquement et techniquement. Dans ces cas, le·la réalisateur·ice et l'équipe technique pourraient souhaiter garder le plus de son direct possible car c'est ce qui sonne plus vrai ; il y a donc un travail de montage conséquent derrière. La post-synchro ne serait pas un remplacement figé, mais un ajout de matière qui propose plus de choix pour s'assurer à la fois de la compréhension et de la véracité de la scène.

Il y a évidemment des cas où d'autres variables jouent dans cette décision : un changement d'intention de jeu qui ne fonctionne pas avec l'image, ou bien une portée de voix insuffisante par exemple.

- Les post-synchros, moi j'essaye de les utiliser que quand c'est vraiment indispensable, donc si je peux ne garder qu'une syllabe ou qu'un mot ou qu'un tout petit morceau je le fais. (MI5)
- On finit bien souvent par revenir à un maximum de directs et à garder juste ce qui est nécessaire. (MI3)

c) Une post-synchro réussie

La notion de la post-synchro réussie, ou la bonne post-synchro a été évoquée de nombreuses fois au cours des entretiens. Il peut être difficile de cerner dès l'enregistrement si une post-synchro pourra être utilisée, notamment à cause des aspects techniques et de la préférence changeante du·de la réalisateur·ice par rapport au jeu.

- La recherche aboutie en séance d'enregistrement

La prise de son de post-synchro constitue la partie la plus importante dans l'intégration d'une post-synchro, car la matière sonore est née en séance d'enregistrement. Il est primordial d'aller jusqu'au bout du processus pour obtenir un résultat crédible et sincère, en veillant à la synchronisation des mouvements des lèvres et des expressions pour transmettre la véritable intention émotionnelle de la scène.

Il est également crucial de ne pas céder à la tentation de valider rapidement des enregistrements sans être suffisamment précis dans le jeu, car cela peut compromettre la qualité globale du projet. Cela peut être difficile avec les plannings serrés des séances de post-synchronisation, les relations humaines complexes et la fatigue accumulée de l'équipe présente, surtout en fin de journée. Cependant, une prise de son non-satisfaisante peut provoquer un travail supplémentaire conséquent à l'étape du mixage. Ce problème n'est souvent pas sous le contrôle de l'équipe technique, mais des temps réduits alloués à la post-synchro, de part des plannings difficiles à établir, font qu'il y a peu de flexibilité.

- Il faut aller jusqu'au bout pour qu'il y ait quelque chose de très crédible et synchronisé, et qu'on ait produit ce qui doit passer au niveau de la sensibilité de la scène. (P1)
- Un des problèmes, c'est qu'on a tendance à valider un peu trop vite des choses qu'on enregistre, pas parce qu'on a envie d'aller vite mais parce qu'on n'est pas forcément assez précis dans le jeu, parce qu'on pense aussi au planning. (MI2)
- Une post-synchro réussie c'est quand tout le monde est content et il n'y a pas de non-dit. (P1)

- Une post-synchro réussie c'est en premier lieu celle qui convainc le réalisateur, notamment en termes de mise en scène et de comédie. (MI3)

L'objectif pendant une séance est de créer une synchronisation qui sonne juste et ne nécessite que peu de corrections supplémentaires. La qualité du matériel fourni facilite grandement le processus de mixage.

- Une post-synchro réussie c'est quand tu n'as presque pas besoin de la traiter au mix, ça sonne juste tout de suite en fait... (DP1)
- Les paramètres de projection, de synchronisme et tout simplement de crédibilité de la phrase au sein du reste... c'est ça qui va être le plus important. Si déjà tout ça est rempli, ça me semble facile à mixer. (MI4)

- L'aspect technique

Dans le cas où il s'agit de faire une post-synchro pour des rattraper l'intelligibilité ou des soucis techniques dans le son direct, la post-synchronisation réussie peut être caractérisée par une parfaite équivalence avec le jeu et l'intention du son direct, en reproduisant fidèlement tous les aspects du jeu, du mouvement du corps et de la respiration. Elle fournit un remplacement technique irréprochable du son direct, tout en préservant son intégrité.

Le son remplacé occasionne un changement de qualité sonore, mais il s'agit de réduire au plus possible la sensation de rupture, à la fois en prise de son, au jeu et à l'étape de mix. C'est une recherche de transparence. Il est essentiel d'apporter suffisamment de corps et de profondeur au son, afin d'éviter un son étriqué ou manquant de présence. Il faut donc faire attention à l'équilibre entre la perche et la HF dans le son d'origine.

Cela dit, l'objectif ultime n'est pas de chercher un raccord purement technique, mais de parvenir à un résultat supérieur à celui obtenu lors du tournage initial dans l'aspect technique, mais également dans le jeu si possible.

- Une post-synchro réussie, c'est une post-synchro qui va être parfaitement égale au jeu du son direct, à la voix du son direct, à l'intention du son direct. Tous les aspects du jeu

et de mouvement de corps de respiration sont parfaitement identiques au son direct, et c'est aussi vraisemblable. (MI1)

- Elle est ratée si on se dit « pourquoi le son il a changé ? (MI1)
- C'est à nous de faire en sorte que ce soit le plus intégré possible, dans l'esprit du direct, mais il y a aussi la véracité du jeu qui dépend du comédien. (MI3)
- Il faut du corps dans le son, parce que dès qu'on est un peu avec un son étriqué où justement il n'y a pas assez de grave et le son change complètement. (DP1)
- Une post-synchro réussie...pour moi c'est faire mieux que ce qu'on avait sur le tournage. (MP1)

- Le respect du sens du film

Il est important de noter que la réussite technique ne doit pas se faire au détriment de la vérité et de la cohérence avec l'ensemble du film. Comme tous les éléments, la post-synchro est un outil qui doit apporter une valeur ajoutée à une scène en répondant à la demande du·de la réalisateur·ice.

Bien que la satisfaction de toutes puisse être un bon indicateur, il est parfois nécessaire de prendre du recul pour évaluer si une autre prise que celle initialement choisie (qui peut être la plus réussie techniquement) serait plus cohérent et bénéfique pour le film. Il peut donc arriver que le jeu d'une post-synchronisation plaise à le·la réalisateur·ice mais soit difficile à intégrer, en raison d'une discordance avec le jeu original. Il peut y avoir aussi le cas contraire, où le raccord semble marcher mais la post-synchro enlève des qualités présentes dans le direct. Il faut alors disposer de plusieurs options qui fonctionnent techniquement, avec la possibilité et la flexibilité de pouvoir choisir celle qui correspond le mieux aux attentes du·de la réalisateur·ice qui peuvent évoluer au cours de la post-production du film.

- Si ce sont des post-synchros faites pour augmenter une émotion ou améliorer le jeu du comédien, dans ce cas-là, celle qui est réussie, c'est celle qui va donner l'émotion et l'intensité qui n'a pas été là au moment du tournage. (MI1)
- Il m'est arrivé à contrario que le jeu d'une post-synchro plaise à un réalisateur et que ce soit une galère à raccorder, parce que peut-être que c'est justement cette bizarrerie pas raccord tout d'un coup qui plaisait aux réalisateurs. Ça s'avère difficile à intégrer mais après c'est une question de parti pris. (MI3)
- Il faut capturer ce qui doit passer au niveau de la sensibilité de la scène. (P1)
- Tout le monde peut être satisfait du résultat sur le coup et quinze jours plus tard avec du recul, on peut retourner vers le direct ou une autre prise qui est peut-être plus dans la continuité du film et fonctionnera mieux. (P1)
- Même si on sent que techniquement on a mieux, si elle permet au film de marcher c'est une super synchro. Le vrai enjeu c'est : est-ce que ça apporte quelque chose au film ? (MI5)
- Techniquement le truc peut être très bien intégré, mais on a perdu une véracité. (MI5)

En parlant de post-synchro réussie, les professionnel·le·s évoquent un certain nombre de fois le raccord technique, qui serait une continuation dans le timbre, couleur et acoustique du son direct. Cependant iels touchent également à la question de la crédibilité, véracité ou de la vraisemblance, qui a été conçue séparément dans le questionnaire mais a été systématiquement abordée en parlant de la post-synchro réussie.

d) La vraisemblance recherchée en post-synchro

- Le lien avec le son direct

La plupart des professionnels ont défini la vraisemblance à partir du son direct. Une bonne post-synchro est celle qui est invisible (correspondance labiale parfaite) et qui ne s'entend pas. C'est donc une réplique que l'on imagine être du son direct.

- Une post-synchro qui ne se voit pas et qui ne s'entend pas, évidemment, en montage et mixage est vraisemblable. (DP2)
- Une post-synchro qui est bien intégrée, on croit que c'est du direct. (MI4)
- La vraisemblance [...] c'est en rapport avec le son direct [...] (DP1)
- Dans le cas le plus simple c'est être raccord avec ce qui précède et ce qui suit. (MI2)

Le lien avec du son direct paraît évident, mais en allant plus loin dans la manière de travailler afin d'obtenir ce résultat, définir ce qui paraît « être du direct » ou « être vraisemblable » se complique infiniment. Cette complexité est due à la fois à nos sensibilités et au fait que la vraisemblance sonore au cinéma est une construction dans un cadre qui évolue, comme nous l'avons vu dans une partie précédente. Nous allons nous concentrer sur la vraisemblance de la voix dans le contexte cinématographique.

Le premier paramètre, et le plus évident, est celui du synchronisme labial. En effet, sans une concordance parfaite entre l'image et le son comme dans du son direct, l'illusion de vraisemblance n'existe plus. Or il y a également des post-synchros que l'on ne voit pas à l'image tel que les répliques rajoutées en voix off, et la vraisemblance qui leur est associée ne repose pas sur le synchronisme.

- Elle doit être synchrone vis-à-vis de du personnage à l'image et des de ces labiales. (MP1)

- Tu peux ne pas avoir de problème de raccord en fait, tu peux avoir juste une phrase toute seule, d'ailleurs une post-synchro ça peut être une voix off. (MI3)
- Quelque chose de vivant, de proche du direct, peut ne pas être vraisemblant. (MI1)

- L'importance du jeu d'acteur et l'aide au jeu

C'est le paramètre qui a réuni unanimement toutes les personnes interviewé.e.s pour ce mémoire. Comme mentionné précédemment dans l'analyse de ce qui constitue une bonne post-synchro, la prise de son en auditorium est le moment le plus important. En effet, si le jeu d'acteur est faux, le raccord technique pourra marcher, mais le·la spectateur·ice ne sera pas convaincu·e de la vraisemblance de la post-synchro. Il vaut donc mieux un jeu réussi et juste avec une prise de son non-satisfaisante plutôt que l'inverse.

La remise des comédien·ne·s dans l'émotion et le mouvement du moment qui se reflètera dans leur voix est un exercice difficile. Il existe plusieurs techniques, allant de la répétition pour préparer les comédiens à l'ajout d'objets dans leurs bouches pour mimer la situation de tournage.

- Il a déjà vécu le truc, il a déjà joué la scène donc c'est possible de reproduire ça en écoutant bien et en faisant attention. » (P1)
- Alors dans ma trousse il y a du sérum physiologique, il y a du chewing-gum pour parler en mâchant, il y a des petits cure-dents pour parler avec un objet dans la bouche [...] j'ai des masques, ça dépend des films, ces objets peuvent aider au jeu en approchant le comédien de son expérience du tournage. (P1)

L'intérêt est de mettre le·la comédien·ne à l'aise et en confiance en couvrant d'abord les séquences plus légères et faciles en post-synchro. Il est notamment souvent plus facile de commencer par des ajouts de texte et voix off (car il n'y a pas de référence à suivre), des phrases courtes et des présences faciles, pour ensuite aller vers des textes plus longs et des émotions fortes. Pour ce qui est des respirations, trouver le bon rythme et la bonne intensité peut s'avérer difficile et prendre du temps.

Dans ce cas où il s'agit de changer d'intention de jeu par rapport au son direct, le plus important est de raccorder le jeu avec les répliques qui entourent la phrase post-synchronisée.

- On aura beau avoir tous les outils les plus performants qui existe, rien n'y fera [...] pour qu'une post-synchro soit réussie il faut vraiment que la voix raccorde. (MI1)
- Si ta prise de son est pas pourrie, techniquement bonne et vraisemblable comme tu dis...donc quand je dis raccord moi vraiment le principal c'est le jeu. (MI2)
- Si le jeu est juste, on accepte beaucoup de choses je trouve. (MP1)

- Simulation des conditions réelles : l'aléatoire et l'accident

La prise de son sur un tournage est faite dans des conditions imposées par le cadre, la lumière et le mouvement de le· la perchman ou perchwoman, ainsi que celui des comédien·ne·s. La simulation des conditions réelles, ou des conditions de tournage, implique la présence de ces acteur·ice·s dans un cadre similaire. Il faut s'assurer de la présence de complexité de la prise de son afin de percevoir une « vie » dans le son. La vraisemblance peut être perçue dans la cohérence de la voix avec le lieu et avec la situation (par ex. la voix de quelqu'un qui court ne peut pas être stable). Le contexte dans lequel le son est émis est alors primordial.

- La vraisemblance ce n'est pas forcément les défauts mais c'est les circonstances qui font qu'il y a des éléments et des parasites. (P1)
- Ce n'est pas forcément une question de raccord avec avant et après, parce que ça peut être la première phrase du film ou une phrase toute seule. Ça peut être tout ce qui est lié au contexte qui fait que c'est vraisemblable. (DP2)
- On a tendance à faire des choses trop statiques [...] faut trouver en fait chaque petite chose qui fait que ça va être raccord, donc c'est la projection de l'énergie comme il faut...c'est bouger. (MI2)

Il y a plusieurs paramètres qui entrent en jeu pour avoir un son vivant avec une part d'aléatoire. On peut avoir une volonté d'avoir un son « propre » en post-synchro, en voulant maîtriser entièrement la voix, en restant plutôt statique et en évitant le risque de la dégradation du signal sonore par la mise en mouvement de la perche et des acteur·ice·s. Une des questions que les professionnels se posent en séance de post-synchro est celle de la limite de mouvement à reproduire en enregistrement. Dans le meilleur des cas, le·la mixeur·se des enregistrements de post-synchro en aura discuté avec le·la mixeur·se du film (si ce n'est pas la même personne). Le·la mixeur·se du film peut vouloir privilégier une prise de son plus statique avec la volonté de recréer tout le mouvement à l'étape du mix, ou au contraire demander un grand nombre de mouvements dès la prise de son.

Le·la perchman ou perchwomen de post-synchro a un rôle primordial et souvent assez peu exploré faute de temps. La position du micro par rapport aux comédiens, son maniement et les détimbrages peuvent pousser très loin l'impression de vraisemblance.

Il est également primordial que les comédien·ne·s puissent bouger dans le cas où iels post-synchronisent des répliques où iels sont en mouvement. En simulant la course, la marche ou même une position corporelle similaire à celle dans l'image, il y a plus de chances de retrouver une texture de voix raccord tout simplement car les cordes vocales ne s'exercent pas de la même manière.

- Il faut avoir l'imperfection du son de tournage où t'as une façon de prendre le son imposée par le cadre et les lumières etc. (P1)
- Parfois c'est très insolite, parce qu'en post-synchro je peux me retrouver à côté du pied du comédien, s'il était en haut d'une roche par exemple, et le son a une couleur bizarre mais ça va marcher. (P1)
- C'est toujours une quête de rajouter une part d'aléatoire, une part d'accident, une part de vivant à un son qui a été enregistré dans une acoustique un peu morte avec peu de mouvements du comédien et du micro. (MI1)
- Ce qui les rend humains et ce qui me touche, ce sont ces imperfections-là. Si c'est trop lisse, ça ne m'intéresse pas. (MP1)

- On emploie le terme « vie », on dit qu'un direct est plus « vivant ». C'est à la fois notre épine dans le pied et le graal d'arriver à faire en sorte qu'une post-synchro soit aussi vivante, aussi pleine de défauts mais dans le sens positif du terme. (MI3)
 - Au-delà du raccord technique : la vraisemblance, une sensibilité humaine

Nous avons déjà évoqué la dualité entre le raccord technique et la vraisemblance dans cet écrit. Le raccord technique est plus facile à définir, car il s'agit d'une dextérité avec différents outils.

Dans une séance de post-synchro, le raccord qui est le plus repris et rappelé aux comédiens est celui de la portée vocale. Il peut être difficile de jauger la portée de la voix quand il s'agit, par exemple, de scènes où les comédiens sont entourés de sons forts dans le film (ex. dans un bar bruyant), contrairement à l'auditorium silencieux. On peut alors fournir un casque aux comédiens afin de provoquer la portée nécessaire.

Le raccord de la hauteur, du timbre et de la couleur de voix sont jugés par les oreilles de toutes les personnes présentes en auditorium ; la direction d'acteur est faite avec ces paramètres en tête. Nous avons parcouru dans une partie précédente les outils qui permettent de réaliser un raccord de l'acoustique (la réverbération), du pitch, et de l'égalisation.

Au-delà de ces paramètres auditifs, la vraisemblance touche à une sensibilité humaine tout autre, qui nous convainc ou non de la vérité du son que l'on entend. Ces paramètres de persuasion sont difficiles à qualifier avec des termes sonores. Il est important de noter qu'il y a plusieurs jeux possibles au son pour une image, et donc plusieurs vraisemblances possibles.

- Est-ce que le raccord et la vraisemblance c'est la même chose [...] si on accepte l'idée qu'il y a plusieurs vraisemblances, oui. On peut très bien changer d'intonation de la même phrase et que ça reste vraisemblable. (DP2)
- Quand j'écoute un comédien, j'ai écouté le direct avant et j'ai une musique dans la tête. Et parfois le comédien propose autre chose, ou le réalisateur demande autre chose et ça me choque, et puis en revoyant ça marche parce que c'est un jeu possible. (DP2)

- Même quelque chose de proche du direct peut ne pas être vraisemblable. Un raccord parfait peut techniquement être complètement viable [...] mais il peut manquer cette espèce de supplément d'âme. (MI3)
- Une infinité de paramètres à auxquels l'oreille est sensible...(MI4)
- On ne va pas l'entendre mais on va se sentir...le but c'est qu'effectivement on ne se dise pas « tiens y a une super bruitage » ou « tiens ce post-synchro est bien. » (MP1)

Une post-synchro qui n'est pas faite pour des raisons techniques existe pour changer le jeu d'acteur. Il est pratiquement impossible d'obtenir une parfaite concordance entre le jeu des acteur·ice·s à l'image et celui enregistré séparément, qui diffère de ce qui a été réalisé sur le plateau. Par conséquent, il est possible de remarquer un décalage entre le jeu des acteur·ice·s à l'image et le jeu tel qu'il est rendu par le son. Or si un jeu qui ne semble pas entièrement raccord avec ce qui l'entoure apporte une intention qui fonctionne dans le cadre du film, il peut être vraisemblable pour les spectateur·ice·s grâce à son apport émotionnel ou artistique.

- Justement c'est génial, en fait au cinéma le but ce n'est pas d'être vraisemblable. (P1)
- Faut arrêter un peu ces a priori de temps en temps et se dire que ça ne va pas être la même perception mais si c'est ce que veut faire percevoir le réalisateur ou la réalisatrice, c'est bien. (DP2)
- C'est lié aussi au réalisme. Nous, techniciens, on a envie d'être réaliste... « ils sont en mouvement, ils sont dehors, il faut timbrer d'une certaine façon » oui, mais le réalisateur ou la réalisatrice peut avoir envie de faire passer une douceur, une intimité et que ça soit ça la priorité. (DP2)

e) Conclusion

Nous avons dégagé différents paramètres qui aident à améliorer la vraisemblance d'un dialogue post-synchronisé : le jeu d'acteur, la vivacité introduit dans le timbre et dans la couleur de la voix par des mouvements aléatoire, l'acoustique et le contexte du son direct. Dans cette optique, nous avons développé un test perceptif visant à évaluer l'importance de ces paramètres, ainsi que l'apport des logiciels actuellement utilisés, dans l'intégration de la post-synchronisation, en ce qui concerne l'impact qu'ils ont sur la perception de la vraisemblance par les spectateurs.

Partie 4 : PPM – Comparaison/Apport des logiciels pour l'intégration des post-synchro

I- Exposé du projet

a) Présentation générale

Nous avons choisi de mener un test perceptif dans le but d'évaluer l'apport des logiciels d'intégration de post-synchronisation et des paramètres sur lesquels ils agissent sur la vraisemblance d'une scène. Nous définissons la vraisemblance comme un rapprochement et une continuité technique du son direct. Ceci inclut l'acoustique, la vie et le mouvement de la voix par rapport au micro ainsi qu'un jeu d'acteur avec un bon synchronisme labial et une bonne correspondance dans l'énergie perçue. Nous avons choisi deux extraits d'environ une minute, afin d'avoir une durée de stimuli pas trop longue et impliquant de la post-synchronisation qui sera mixée de différentes manières à l'aide de divers logiciels. L'objectif sera donc de déterminer l'impact de certains paramètres ou logiciels sur la perception du public en ce qui concerne la vraisemblance des dialogues présentés.

Ce test se déroulera en 2 parties, qui se répéteront 2 fois pour 2 séquences différentes, soit 8 écoutes au total.

Partie 1 :

Le spectateur·ice regardera 2 extraits : le son direct mixé ainsi qu'une version post-synchronisée et intégrée. Il notera alors sa préférence pour un des extraits. Le jeu d'acteur étant différent dans les deux cas, il nous a semblé plus intéressant de séparer cette écoute de la comparaison entre les différents mix des post-synchros. Cette étape nous permet de voir si les spectateur·ice·s détectent la présence d'une post-synchro intégrée, et s'ils ont des préférences autres que techniques (jeu d'acteur, compréhension, etc.). On répète l'expérience pour les deux séquences.

Partie 2 :

Une série d'écoutes sans différence d'interprétation du jeu aura lieu. Le même extrait (post-synchronisé) sera diffusé avec 3 mixages différents, traités avec les logiciels suivants :

Mix A : Chameleon + Boom-It + EQ Match

Mix B : Boom-It + EQ Match

Mix C : Chameleon + EQ Match

Cette deuxième étape permettra de voir comment les paramètres sur lesquels jouent ces logiciels, notamment l'acoustique et la vivacité de la voix, influencent la perception du·de la spectateur·ice par rapport à la vraisemblance d'une scène.

- L'évaluation de l'acoustique : il s'agit d'ajouter et de raccorder la réverbération de la voix post-synchronisée à celle de la voix dans le son direct.
- L'évaluation de la vivacité de la voix : ce paramètre correspond aux changements de timbre et de couleur dans la voix apportée par les mouvements des comédiens ainsi que par les mouvements de perche dans un son direct.

Cette étape permettra également de constater l'apport de Boom-It et de Chameleon dans le cadre de l'intégration des post-synchro.

On répète l'expérience pour chacune des deux séquences.

Le test aura lieu en auditorium de mixage, ce qui permet d'avoir la main sur l'ordre des extraits plus facilement et de privilégier un passage aléatoire des extraits aux différents groupes qui passeront et de garantir une qualité de l'écoute. Afin d'éviter le plus possible le biais d'ordre, nous allons alterner l'ordre des écoutes pour chaque groupe (constitué de 2 à 3 personnes). Le questionnaire tel qu'il a été présenté aux spectateur·ice·s pendant le test est présenté ici en annexe.

b) Choix des extraits

Nous avons choisi de travailler sur des extraits de film déjà existants, car il n'était pas possible, dans le temps et budget imparti, de tourner et post-synchroniser des extraits à la hauteur de la qualité désirée (à la fois techniquement et en termes de jeu).

Quatre extraits de film nous ont été confiés par la directrice de plateau Anne Maisonhaute, qui nous a aidé·e·s à les choisir soigneusement afin d'illustrer différents cas

d'intégration de post-synchro. Les deux premiers extraits proviennent du même film et sont presque entièrement post-synchronisés. Le premier se passe dans un lieu réverbérant avec un orateur et une foule excitée. Le deuxième consiste en des confidences et des voix très proches par rapport à la prise de son. Le troisième se passe dans une petite maison ; il s'agit d'une conversation entre une fille et son père, où les post-synchros sont plus éparés. Le quatrième se passe à l'extérieur, et contient des ajouts de respirations et de présence, mais peu de dialogues.

Nous avons préféré travailler sur les deux premiers extraits car il nous semblait plus intéressant de sélectionner des extraits avec une grande quantité de post-synchros dialoguées afin de faciliter l'écoute des sujets de test, et de mettre en œuvre les logiciels sur la quasi-totalité des séquences. Les respirations et présences dans le quatrième extrait étant des ajouts qui n'existent pas dans le son direct, cela induisait un grand biais au test. Ces éléments sont également intégrés en montage son avec un travail de recalage consécutif, ce qui nous semblait être un cas à part ; plusieurs professionnel·le·s nous ont confirmé que c'était un exercice particulier, même au sein de l'intégration des post-synchros.

Nous avons alors récupéré plusieurs éléments pour chaque extrait :

- La vidéo de référence
- La session ProTools avec des enregistrements de post-synchro triés
- Une piste de son témoin avec le montage des sons directs dans l'état avant la post-synchro
- Un bouchage : une piste contenant les paroles qui n'ont pas besoin de post-synchro et une ambiance sonore continue pour boucher les moments de trou (moments de paroles à remplacer)
- Le chutier de post-synchro : des pistes contenant d'autres prises de post-synchro au cas où l'on veut revenir sur d'autres choix
- La musique dans le cas où elle est présente

La première étape a ensuite été de nettoyer les sons directs et de recalibrer les post-synchros pour une écoute confortable et un meilleur synchronisme labial. Nous avons également ajouté quelques ambiances et bruitages pour se rapprocher d'une bande son de film. Il ne nous était pas possible, dans le temps imparti et avec nos compétences, de pousser très

loin chaque étape de la post-production sonore, ces traitements étaient réalisés surtout pour assurer une écoute sans biais majeurs.

c) **Choix et utilisation des logiciels d'intégration**

Comme nous avons pu le constater, les post-synchros sont intégrées à l'aide plusieurs types de traitement, correspondant à plusieurs plug-ins (EQ, compression reverbe, pitch, etc.). Nous avons choisi de nous concentrer sur deux éléments principaux : l'acoustique et le détimbrage de la voix par rapport aux mouvements perçus à l'image. Pour évaluer l'impact d'autres logiciels agissant sur des critères différents, il aurait fallu réaliser des tests plus élaborés et complexes, ce qui n'était pas possible dans le temps imparti. De plus, dans le cadre de notre sujet, il est plus intéressant d'évaluer les nouveaux logiciels sur le marché, tels que Chameleon et Boom-It, qui ont été développés pour répondre aux besoins de l'intégration des post-synchros, et voient une utilisation accrue dans ce contexte précisément.

- EQ Match d'iZotope

Nous avons choisi d'employer ce logiciel afin d'harmoniser les couleurs des voix de la prise de son direct à celle de la post-synchro. Nous avons jugé essentiel d'inclure cette étape afin d'éliminer les variations significatives potentielles entre les couleurs de voix, qui auraient pu fausser l'évaluation axée sur les deux autres logiciels. En effet, la différence perçue aurait pu prédominer sur les effets perçus de Boom-It. Nous avons décidé de ne pas évaluer ce logiciel car il fait partie de la méthodologie de travail courante aujourd'hui et ne présente pas d'innovations particulières.

- Chameleon

Logiciel présent depuis 2021, Chameleon a rapidement transformé le travail d'intégration des post-synchro grâce à son ergonomie et sa précision. Il est de loin le logiciel de création de réverbération raccord préféré des mixeur·se·s qui l'ont déjà expérimenté, d'après nos discussions avec les professionnel·le·s. Nous l'avons choisi pour évaluer l'apport du réacoustiquage sur la vraisemblance des post-synchros dans deux extraits qui présentant des caractéristiques de réverbération différentes.

L'utilisation de Chameleon s'est avérée très intuitive. Une fois qu'un son de référence lui est fourni, le logiciel génère une réverbération qui fonctionne comme une réverbération à convolution. Nous avons agi sur les paramètres proposés (voir p.51) pour essayer d'obtenir l'acoustique la plus proche possible de celle présente dans le son direct. Nous n'avons appliqué aucune correction d'EQ sur le résultat, car nous voulions justement évaluer sa capacité à raccorder les différents paramètres de l'acoustique, y compris la couleur de la pièce. Nous avons écouté et comparé plusieurs moments du son direct afin de choisir les répliques à lui fournir pour calculer la reverbe. Les voix à raccorder étant globalement au même plan, respectivement pour chaque extrait, nous avons remarqué que le choix des répliques n'influçait pas grandement la réverbération générée. Nous n'avons pas rencontré de difficultés à cette étape ; le logiciel a généré sans exception des réverbérations raccords à chaque fois que cela lui était demandé.

- Boom-It

Il s'agit d'un logiciel et un concept nouveau pour ajouter du mouvement et de l'aléatoire dans un son plus statique. Il existe évidemment des films en France où il n'est pas possible de faire appel à un-e perchman ou perchwoman pour l'enregistrement des post-synchros faute de budget, et l'utilisation de ce logiciel est d'autant plus d'intéressant. Cependant, nous souhaitons évaluer l'apport de Boom-It dans le cas où l'on accorde déjà une attention particulière à ce détail lors de l'enregistrement de post-synchros afin d'éviter des voix statiques.

L'interface de Boom-It nous était déjà familière grâce à nos recherches préalables, mais le concept d'utilisation étant nouveau, cela nous a demandé plus de réflexion avant de pouvoir agir sur les paramètres de la bonne manière afin d'obtenir le résultat souhaité. Nous n'avons pas utilisé le module de changement de couleur de microphone car, comme on a vu, il est destiné à des prises de sons faites avec des micros de studio, ce qui n'était pas notre cas. De plus il aurait changé l'égalisation du son, chose que l'on ne souhaitait pas, car dans ce cas il aurait été difficile de séparer l'apport d'EQ Match et de Boom-It sur la couleur du son.

Pour le module *Head Tilt and Angle*, nous avons commencé par reproduire de manière exacte le mouvement de tête des comédien·ne·s tels qu'ils sont perçus à l'image. En apparence ceci est un exercice simple, mais la question du placement du micro s'est posée assez rapidement. En effet, si l'on voulait reproduire exactement les mouvements, il fallait avoir une

référence : le placement et les mouvements de perche de la prise de son en post-synchro. Il est difficile de se rendre compte de ces détails simplement en écoutant les prises, surtout à notre niveau d'expérience. Nous avons alors décidé de comparer les prises de son direct avec les post-synchro afin d'écouter les différences de couleur, pour choisir « axe neutre » selon les cas. Par exemple, dans un cas où deux personnages se font face, nous avons fait les mouvements en commençant avec un axe neutre décalé de 15°, car la couleur des voix nous semblait correspondre le mieux au son direct à cette position du micro. Tous les mouvements de cet exemple se sont faits avec un décalage de 15° selon la position neutre proposée par Boom-It.

Le module liveliness a été réglé entre 15% et 30% selon les cas.

Nous avons réalisé un seul mix avec les trois logiciels cités ; nous avons ensuite « by-passé » les plug-ins non nécessaires suivant les cas lors des exports.

Nous avons décidé d'éliminer *Dialogue Match* pour plusieurs raisons. Tout d'abord, il partage une fonctionnalité commune avec Chameleon : le module utilisé pour raccorder des reverbes. Cependant, lors de nos tests de comparaison, ce module nous a paru moins convaincant que celui de Chameleon. Si nous avions inclus ce logiciel, nous aurions dû comparer ces deux outils entre eux, ce qui n'était pas l'objectif principal de cette PPM. De plus, une ambiance raccord nous avait déjà été fournie dans la session de montage des directs des extraits choisis ; nous n'avons donc pas besoin du module *Ambiance Match* qui permet de générer un fond d'air raccord avec le son direct. Enfin, l'inclusion de *Dialogue Match* aurait significativement prolongé la durée des tests perceptifs. Nous nous sommes donc limités à Chameleon et Boom-It afin d'identifier quels paramètres influencent la perception de la vraisemblance et comment ces logiciels affectent ces paramètres.

II- Présentation et analyse des résultats

Dans cette étude nous avons eu 40 sujets, dont 26 qui ont de l'expérience dans le domaine du son et 14 n'en ont pas. Voici la répartition des tranches d'âge de ces participants :

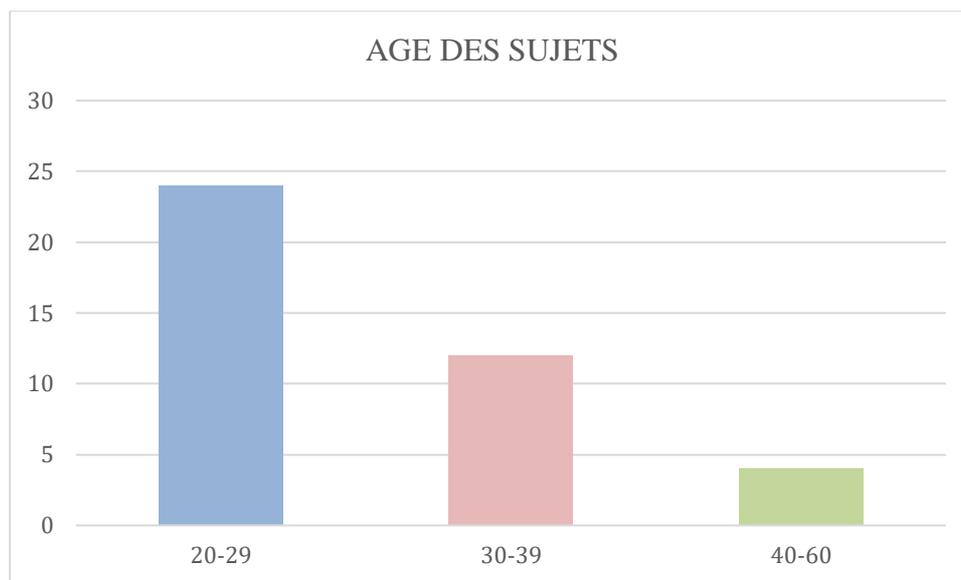


Fig 4.2.1 : La répartition d'âge des sujets

a) Partie 1

Dans cette partie, nous avons choisi de distinguer les résultats des individus ayant une expérience en tant qu'étudiant·e, professeur·e ou professionnel·le dans le domaine du son, de ceux qui n'ont pas cette expérience. Les différences observées entre ces deux groupes dans la première partie du test sont assez significatives pour être intéressantes. Par conséquent, nous désignerons les individus possédant une expérience dans le domaine du son comme des "experts", tandis que les autres seront qualifiés de "novices".

Il est intéressant de constater la cinéphilie de chaque groupe avant de commencer l'analyse des résultats.

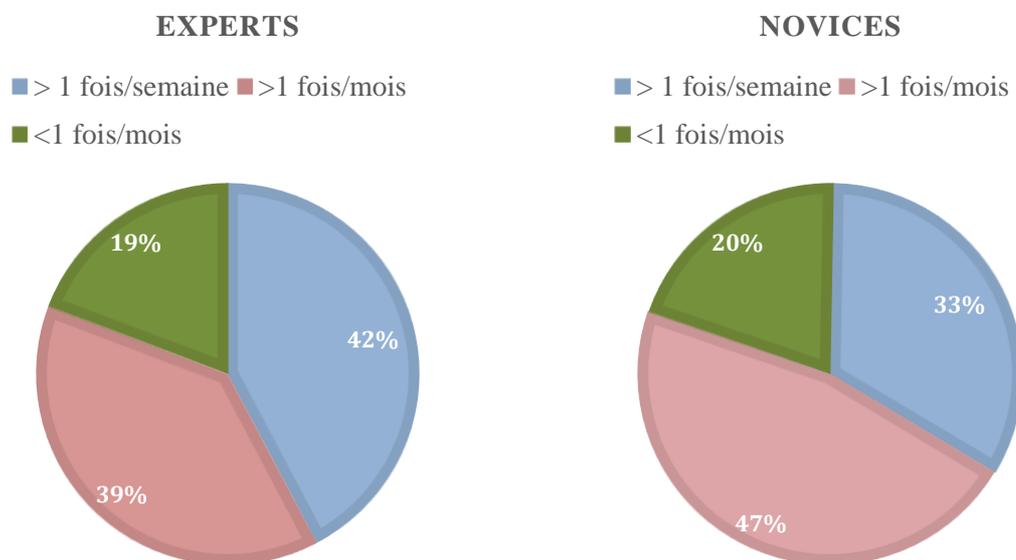


Fig 4.2.2 : Fréquence de visionnage de films en salle par groupe

Il est essentiel de noter à ce stade que, avec un échantillon de moins de 100 participants, nos pourcentages ne sont pas statistiquement rigoureux. Cependant, afin d'obtenir une meilleure vision d'ensemble et d'appréhender les proportions, nous avons opté pour cette présentation qui nous semblait adaptée. Chez les experts, une personne représente environ 3,8% du graphique et chez les novices ce chiffre s'élève à environ 7,1%.

Nous avons des proportions très similaires de la fréquence de visionnage de film en salle pour les deux groupes. A peu près 80% de nos sujets, les deux groupes confondus, sont des spectateur·ice·s (en salle) de films régulièr·e·s, et ont donc une bonne connaissance des codes cinématographiques en France.

Pour la première partie du test, nous avons montré deux versions de deux séquences aux sujets, soit quatre écoutes en tout. L'extrait A de chaque séquence contient du son direct uniquement et l'extrait B est post-synchronisé et mixé au mieux de nos compétences afin de bien l'intégrer. Pour les deux extraits A et B de chaque séquence, nous avons demandé à nos sujets de noter si d'après eux, il s'agissait entièrement de son direct. Nous les avons encouragés à répondre « non » à cette question pour la plus faible perception de la présence de post-synchro.

Est-ce que cet extrait contient exclusivement du son direct ?

(Résultats des deux séquences confondues)

Experts

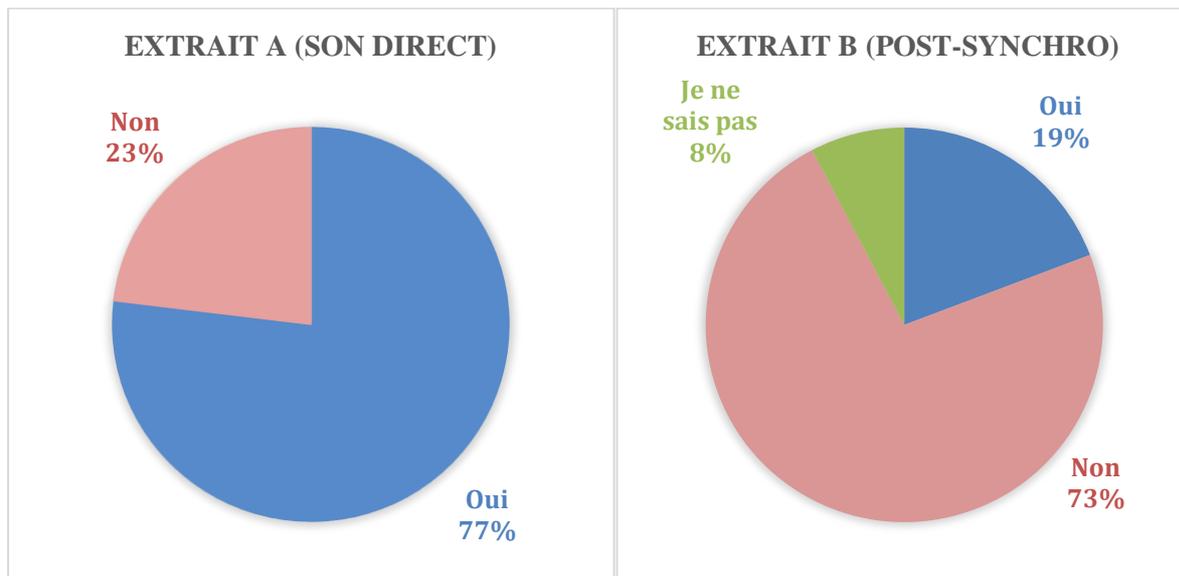


Fig 4.2.3 : Résultats en pourcentage de la question chez les experts

Novices

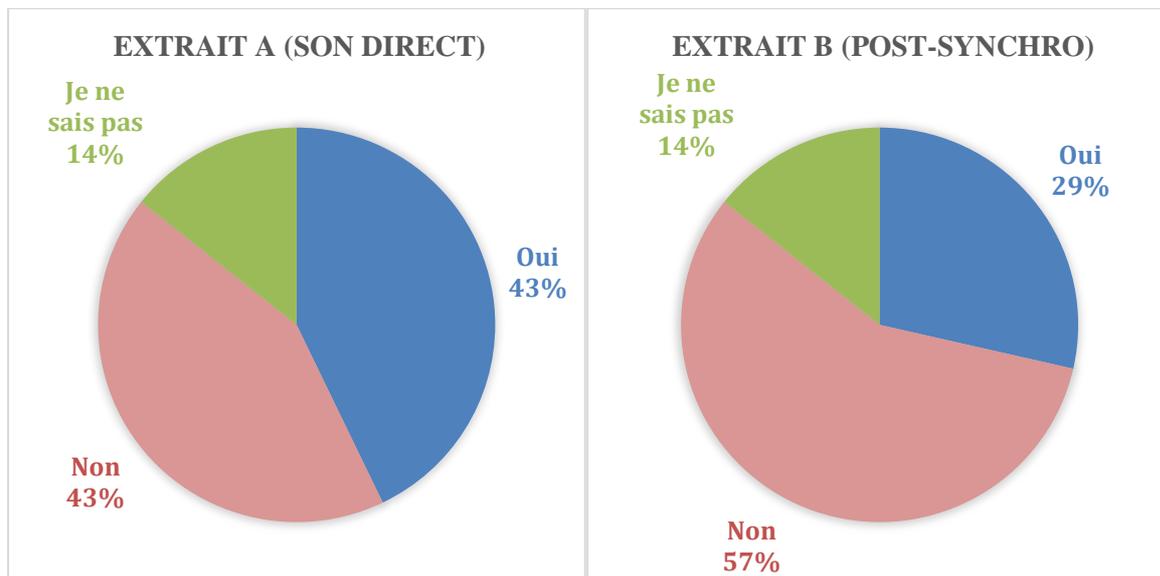


Fig 4.2.4 : Résultats en pourcentage de la question chez les novices

Nous remarquons que les personnes expertes arrivent plus facilement à entendre ou ressentir la présence des post-synchros. 77% (de 26) des experts ont identifié correctement le

son direct de l'extrait A, contrairement à 46% (de 14) des novices. 73% des experts ont identifié correctement la présence de la post-synchro dans l'extrait B contrairement à 57% des novices, la différence est alors moins grande que dans l'extrait A.

Les experts sont également plus assuré·e·s dans leurs choix dans l'extrait A ; proportionnellement plus de personnes novices n'ont pas su s'il s'agissait du son direct ou post-synchronisé. Quant à l'extrait B, 2 personnes de chaque groupe n'ont pas su identifier le direct ou la post-synchro.

Nous avons ensuite demandé aux sujets de noter la vraisemblance des deux extraits sur une échelle de 1 à 5. Nous avons défini la vraisemblance à l'aide des critères évoqués lors de l'analyse : le lien avec le son direct, le jeu d'acteur, le mouvement et la vivacité de la voix (qui reprend le critère de l'aléatoire) et une sensation globale. Il a été demandé aux sujets de se concentrer surtout sur les voix, l'élément au centre de notre étude.

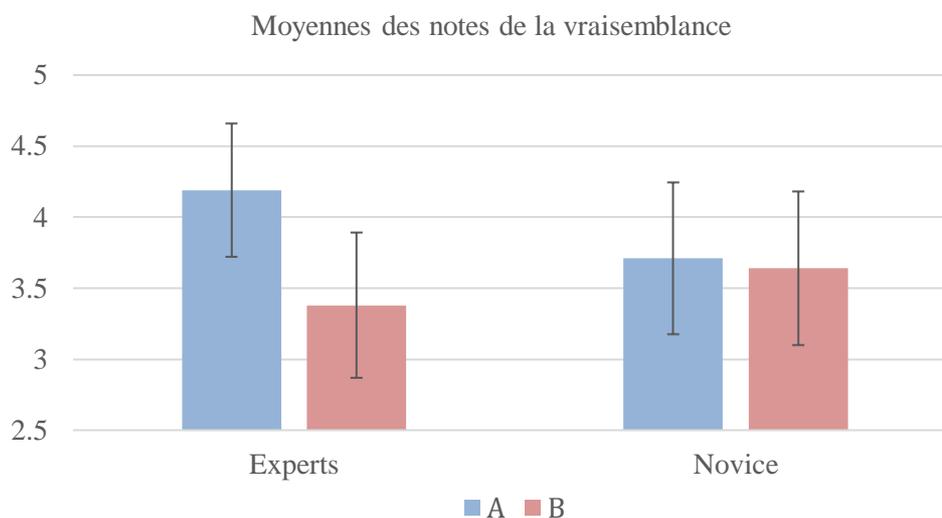


Fig 4.2.5 : Moyenne et écart-types des notes de la vraisemblance par groupe et extrait

Nous avons essayé de nous mieux d'expliquer le concept de la vraisemblance, en utilisant des exemples concrets d'autres extraits que ceux du test si nécessaire. Malgré cela, nous ne pouvons pas être sûr que les novices et les experts aient compris ce concept au même degré, ce qui introduit un biais potentiel. Néanmoins, il n'y avait pas de grande différence dans la manière de noter la vraisemblance entre les deux groupes dans la partie 2 du test (p.90). On considère alors que les différences ne sont pas significatives en termes de stratégie d'écoute ; cela peut nous rassurer par rapport aux résultats de la partie 1 du test.

Il y a une différence très claire dans la perception des deux groupes : les novices remarquent beaucoup moins de différences en termes de vraisemblance que les experts. Cela se reflète également dans la préférence des sujets entre les deux extraits.

Quel extrait avez-vous préféré ?

A : Son direct, B : Post-synchro

Seq 1

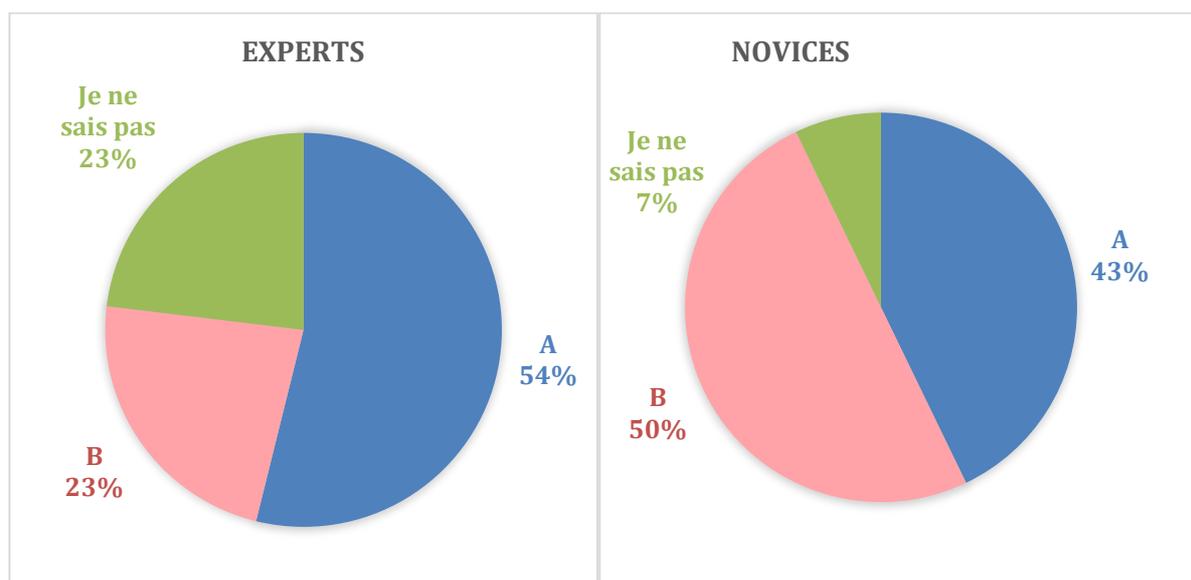


Fig 4.2.6 : Pourcentage de sujets préférant chaque extrait de la séquence 1

Seq 2

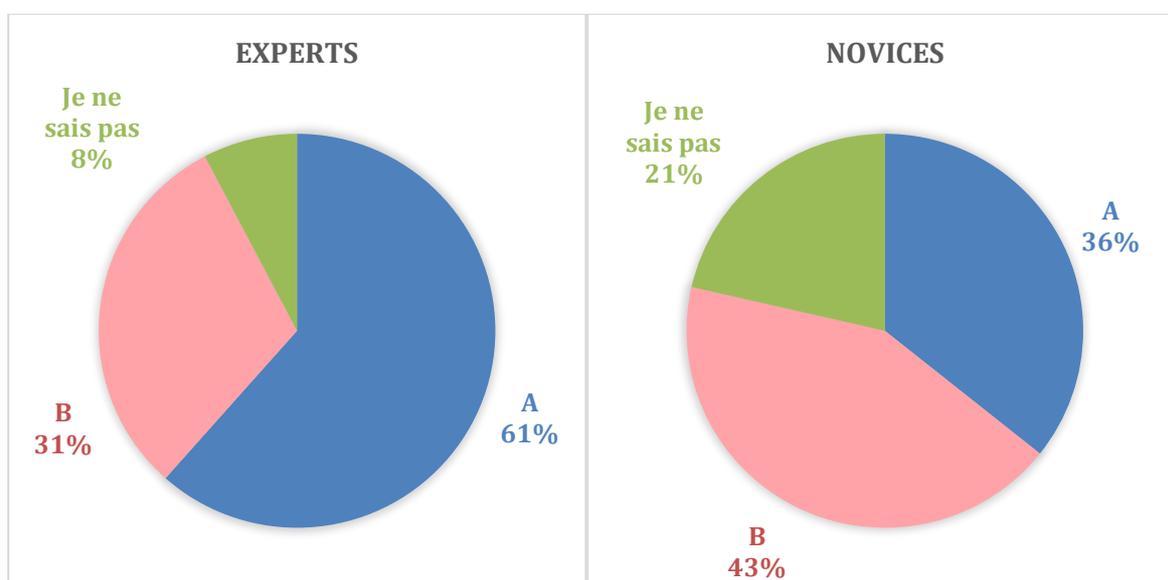


Fig 4.2.7 : Pourcentage de sujets préférant chaque extrait de la séquence 2

Les novices ont une préférence significativement plus élevée des post-synchros par rapport aux experts, la tendance étant nettement plus marquée pour la séquence 1. Nous remarquons en effet une corrélation entre la préférence et l'impression de vraisemblance. Ceci peut faire partie d'un biais introduit par le test qui introduit la notion de la vraisemblance avant de poser la question de la préférence.

Nous avons également demandé aux sujets de citer leurs raisons de préférence pour tel ou tel extrait, pour mieux comprendre leurs choix et critères de jugement.

- Raisons de préférence du son direct (extrait A) :

Séquence 1 :

Chez les experts, 2 d'entre eux ont préféré l'extrait A en raison d'une meilleure synchronisation, 4 en raison de son réalisme, 4 ont préféré le jeu des acteurs, 2 ont noté des présences corporelles plus vivante, 1 personne a remarqué que l'extrait B présentait des sons trop nets et 1 personne a ressenti une gêne dans l'extrait B.

Chez les novices, 3 personnes ont préféré le jeu des acteurs de l'extrait A, 2 personnes ont préféré le son de la voix, et 1 personne a trouvé cet extrait plus crédible et immersif.

Séquence 2 :

Chez les experts, 4 personnes ont préféré l'acoustique des voix dans l'extrait A, 4 ont préféré le jeu des acteurs, 3 ont trouvé les voix et les espaces plus réalistes, 2 ont ressenti une présence corporelle plus marquée, 2 se sont sentis davantage immergés dans le film, et 1 personne a noté un meilleur synchronisme.

Chez les novices, 2 personnes ont trouvé les voix de l'extrait A meilleures, 2 personnes ont simplement répondu en fonction de leur ressenti, et 1 personne a estimé que le jeu des acteurs correspondait mieux à l'image.

- Raisons de préférence de la post-synchro (extrait B) :

Séquence 1 :

Parmi les experts, 3 personnes ont préféré le jeu dans l'extrait B, 2 ont trouvé les voix plus vivantes par rapport au son direct, 1 personne a apprécié l'acoustique de la salle, et 1 personne a noté une continuité entre les plans dans le son et le jeu des acteurs. Une personne n'a pas donné de raison.

Chez les novices, 4 personnes ont trouvé la diction meilleure et plus compréhensible que dans le son direct, 2 personnes ont ressenti une immersion plus facile dans le film, et 1 personne a préféré la qualité sonore globale.

Séquence 2 :

Chez les experts, 2 personnes ont trouvé le jeu des acteurs meilleur dans l'extrait B, 2 ont noté une présence plus vivante, 1 personne a trouvé les voix naturelles, et 1 personne a préféré le timbre des voix.

Parmi les novices, 2 personnes ont trouvé les voix plus naturelles et fidèles, 2 personnes ont préféré la compréhensibilité des répliques, et 1 personne a trouvé le timbre des voix meilleures.

Une personne novice a marqué qu'elle préférerait l'authenticité de l'extrait A mais l'intelligibilité était meilleure dans l'extrait B. Cette remarque est assez représentative des caractéristiques de chaque extrait. En effet, un des intérêts principaux de la post-synchro (et la raison d'existence de la plupart des répliques post-synchronisées dans le cinéma français moderne) est l'amélioration de la compréhension des répliques. Cela était également le cas pour les deux séquences qui ont été présentées aux sujets.

- Conclusion de l'analyse de la Partie 1

Il est intéressant de noter que les experts ont globalement privilégié le caractère naturel du son direct, avec ses imperfections, par rapport à la compréhension des dialogues. Même parmi ceux qui ont préféré l'extrait B chez les experts, aucun-e n'a mentionné l'intelligibilité des dialogues. Cela reflète la tendance en France de préserver à tout prix le son direct et la

valeur qui lui est accordée par ceux qui travaillent dans les domaines du son à l'image et d'autres métiers du cinéma.

En revanche, près de la moitié des novices qui ont préféré la version post-synchronisée ont cité la « compréhension des répliques » comme principal facteur dans leur choix. Ils représentent la majorité du public cinématographique, et malgré leur effectif plus réduit par rapport aux experts dans cette étude (14 contre 26), cela souligne l'importance de cette question d'intelligibilité pour le public. Même avec toutes les recherches visant à ajouter des imperfections pour se rapprocher du son direct, il est primordial de privilégier ce dont le public a besoin, à savoir une compréhension claire des dialogues. Ce point a été abordé par certains professionnel·le·s lors des entretiens. La recherche de cette vraisemblance est d'autant plus complexe et intéressante, car il faut trouver un équilibre entre ces deux critères : la clarté et les accidents aléatoires qui participent à cette sensation d'un « son naturel ».

b) Partie 2

Dans cette partie, nous avons fait écouter trois versions de mix de post-synchros qui ont été intégrées à l'aide de différents logiciels. Les sujets devaient ensuite noter trois critères sur une échelle de 1 à 5 : la vraisemblance, la vivacité et l'acoustique. Pour cette partie, il nous a semblé plus pertinent de grouper ensemble les réponses des personnes avec et sans expérience dans le domaine du son, car il n'y avait pas de différence notable dans leur manière de noter les trois critères.

Les trois extraits pour chacun des deux séquences ont été montrés dans différents ordres afin de contrebalancer un effet d'ordre qui pourrait biaiser les résultats.

Nous avons réalisé plusieurs analyses de variance à un facteur³⁰, calcul que nous désignerons par l'acronyme ANOVA. La différence de moyennes entre les notes pour chaque critère par rapport au logiciel utilisé s'est révélée significative et ce, pour les deux séquences. La valeur de la probabilité p est restée largement inférieure à 0,05 pour chacune des ANOVA.

30 L'analyse de variance à un facteur est un modèle statistique qui permet de déterminer si les moyennes de multiples groupes provenant d'une même population présentent une différence significative. Ce test s'applique lorsqu'on cherche à déterminer l'influence d'une variable explicative catégorielle, aussi appelée variable indépendante, sur la loi de distribution d'une variable continue, dite variable dépendante. Par exemple, il peut être utilisé pour étudier la répartition des notes de la vraisemblance en fonction du logiciel utilisé.

• Séquence 1

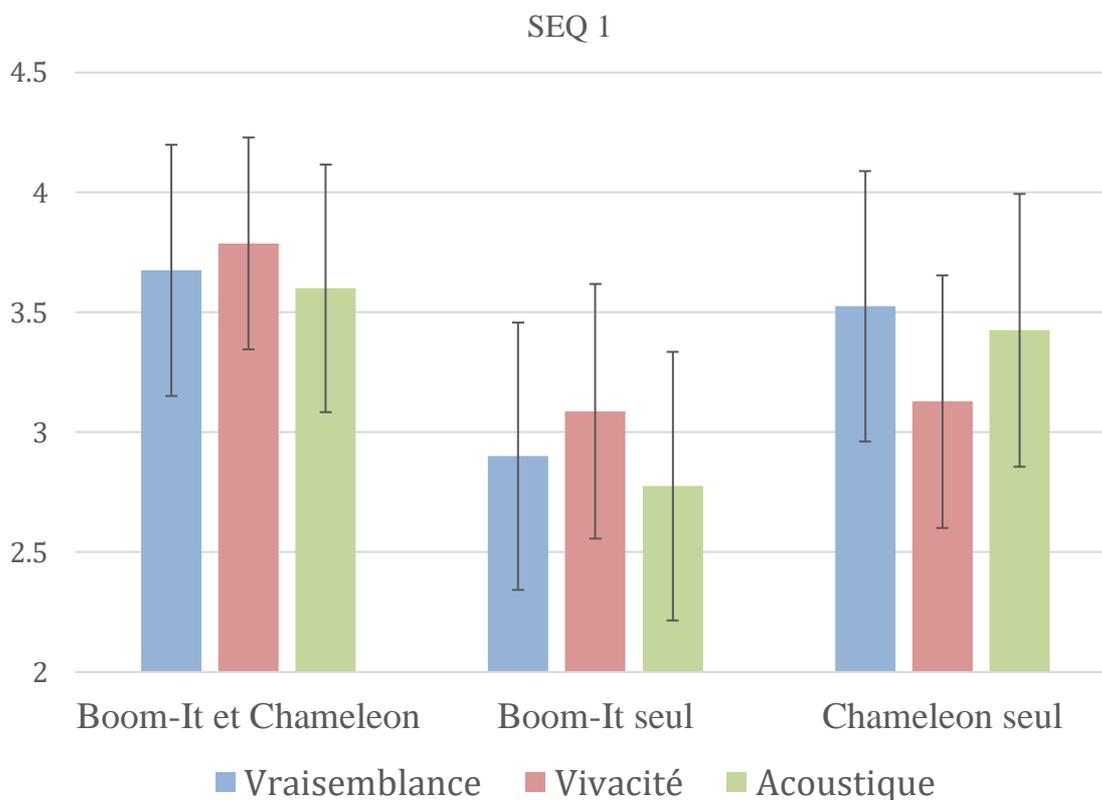


Fig 4.2.8 : Moyenne et écart type des notes par critère et par extrait de la séquence 1

Ce graphique représente la moyenne des notes sur les trois critères à évaluer pour la séquence 1 par rapport à l'utilisation des logiciels.

D'après le graphique, l'absence de Chameleon, et donc de la réverbération est plus ressentie que l'absence de Boom-It (les mouvements de tête et de la perche). L'absence de réverbération est également ressentie comme une absence de vivacité car une voix sans réverbération - donc qui n'a pas été placée dans l'espace du son direct - semble plus « collée ». L'absence de Boom-It semble moins affecter la note de la vraisemblance que l'absence de la reverbe, mais les sujets ont ressenti également une vivacité plus basse. Iels ont noté légèrement plus bas l'acoustique dans ce cas, mais la différence est très minime, de l'ordre de 0,2.

Nous pouvons en conclure que même si l'absence de Boom-It affecte la vivacité dans cet extrait, l'absence de Chameleon a un effet négatif sur les trois critères, et donc est plus ressentie. Il est important de mettre ces résultats en perspective avec la nature de l'extrait : en effet, la séquence 1 se déroule dans un lieu réverbérant où un orateur et une foule excitent

l'acoustique de la pièce. Il est donc naturel que le manque de réverbération sur les enregistrements post-synchronisés soit perçu comme une mauvaise vraisemblance. Cela dit, même dans ce cas où la post-synchronisation a été réalisée avec un perchman mobile, Boom-It apporterait toujours une vivacité perceptible aux voix grâce à son ajout de mouvement. Nous observons une différence notable entre la vivacité en utilisant uniquement Chameleon et en utilisant Chameleon avec Boom-It.

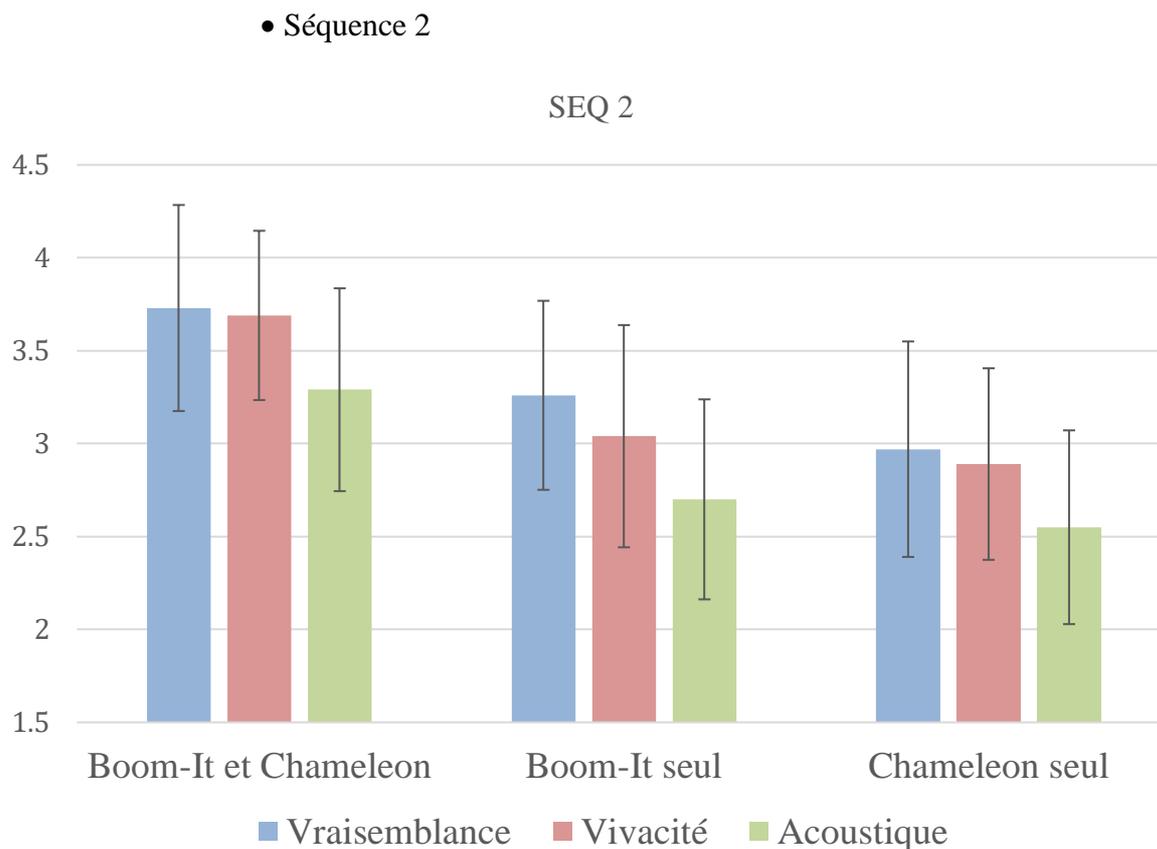


Fig 4.2.9 : Moyenne et écart type des notes par critère et par extrait de la séquence 2

Dans cette séquence, l'absence de Boom-It est plus ressentie que l'absence de Chameleon. On remarque une baisse nette dans les trois critères dans les deux cas. Même si les deux logiciels semblent jouer un rôle plus égalitaire dans la perception des critères par rapport à la séquence 1, les traitements issus de Boom-It semblent apporter plus de sensation de vraisemblance.

Les voix étant très intimistes et proches dans cette séquence, la perte de la réverbération est légèrement moins ressentie que celle des mouvements de tête. En revanche, l'absence de

Boom-It est perçue comme une baisse dans la crédibilité de l'acoustique. Ceci peut être expliqué par le fait que dans ce cas, avec des voix proches et « sèches », les changements de couleurs qu'apportent Boom-It font vivre l'espace autant, voir plus que l'ajout de Chameleon.

- Conclusion de l'analyse de la Partie 2

On en conclut que dans les espaces acoustiquement plus neutres, et surtout avec des voix moins réverbérées, la perception de la couleur de la voix est d'autant plus importante et influe sur la perception de l'espace en plus de la vivacité des voix. Il serait donc plus intéressant d'ajouter du mouvement aux répliques, ce qui semble logique car, à défaut de pouvoir s'intégrer dans l'espace grâce à une réverbération marquée, elles s'intègrent grâce à la vraisemblance de la couleur et des détimbrages.

III- Discussions et limites du test

Cette expérience proposait des termes limités aux sujets pour classer leurs ressentis par rapport à l'intégration des post-synchros. Comme nous avons pu le conclure grâce aux analyses des entretiens, une intégration réussie des post-synchros se repose sur un grand nombre de paramètres. Ce test, avec son vocabulaire limité n'évaluait que trois de ces critères ; or il est difficile de traduire ce que l'on entend en notes pour des paramètres aussi restreints. Plusieurs sujets nous ont exprimé la difficulté de cet exercice.

De plus dans la deuxième partie du test, dans un souci d'éviter des mixages trop caricaturaux et grossiers, nous avons préféré faire de notre mieux avec chaque logiciel afin d'intégrer des post-synchros. Ceci a donné lieu à des différences assez subtiles entre les trois versions, ce qui a ajouté à la difficulté du test.

Ni les participants ni nous n'avions le temps de réaliser des écoutes sur une dizaine d'extraits qui auraient présenté une plus grande variété de situations de post-synchros, ce qui aurait permis d'obtenir des résultats plus représentatifs de la perception de ces différents critères sur lesquelles les logiciels agissent.

Il est aussi important de noter que les participants étaient dans une posture de test d'écoute où ils devaient évaluer des critères précis, ce qui les a probablement mis dans un état

de concentration plus élevée que lorsqu'ils regardent un film pour le divertissement. Ils ont donc porté une attention particulière aux détails, ce qui diffère de l'écoute qu'ils auraient lors d'un visionnage d'un film de deux heures en salle.

Une bonne partie des participant·e·s étant des professeurs ou des étudiant·e·s de l'ENS Louis-Lumière, ou bien des connaissances professionnelles, le public était également assez homogène en termes de cinéphilie. Iels avaient une fréquentation de salle de cinéma assez élevée ; faisant en moyenne au moins 21,67 entrées au cinéma par an. La moyenne française est de 3,8 entrées par spectateur·ice par an (d'après le bilan du CNC de 2022). Or 72% du public de cinéma sont des occasionnels (qui vont au cinéma moins d'une fois par mois), ce qui correspond à 20% de nos sujets. Il aurait été souhaitable d'avoir une représentation plus diversifiée du public cinéphile actuel, ainsi qu'un nombre équivalent de participants experts et novices. Cela aurait permis d'obtenir des résultats plus représentatifs et généralisables à l'ensemble de la population.

CONCLUSION

Contributions

À travers ce mémoire nous avons présenté un état de l'art de la post-synchronisation en France à l'heure actuelle. Les enjeux, méthodes et outils ont été explorés à l'aide des professionnel·le·s travaillant dans le domaine. La question de la vraisemblance est apparue comme un aspect incontournable de la post-synchronisation. La partie pratique s'est alors attachée à évaluer trois critères qui influent sur la perception de l'intégration d'une post-synchro chez le·la spectateur·ice. Les résultats nous apprennent que la couleur et le mouvement de la voix sont essentiels à une bonne intégration, surtout dans des acoustiques mortes. Au-delà des critères de raccord technique, l'introduction de l'aléatoire et de la vivacité dans les prises de son ainsi qu'à l'étape du mix est d'une grande aide pour que le son paraisse vraisemblable. Cependant, il est important de noter que les spectateur·ice·s de cinéma qui ne sont pas familier·e·s avec les subtilités sonores préféreront souvent des dialogues intelligibles, même au détriment d'un réalisme absolu.

L'émergence de l'intelligence artificielle (IA) et de la synthèse vocale soulève des questions sur l'avenir de nos métiers dans l'industrie audiovisuelle. Bien que certaines tâches puissent être automatisées, l'expertise humaine et la sensibilité artistique restent indispensables pour assurer une qualité sonore optimale et une interprétation fidèle des dialogues. Il serait intéressant d'interroger et d'explorer les possibilités de l'IA, et d'appréhender la manière dont elle peut faire évoluer à la fois les logiciels et le workflow. En effet, une automatisation des tâches purement techniques peut donner plus de place à des expérimentations créatives.

Limitations

Dans cet écrit nous avons traité de manière globale les post-synchros effectués dans le cinéma verbo-centré français. Or chaque film a ses propres exigences en matière de post-synchro, dictées par le type de projet et les besoins artistiques. Les films d'action ou d'aventure peuvent nécessiter une concentration sur la portée de la voix pour que les répliques soient compréhensibles et réalistes au-dessus des niveaux sonores élevés, tandis que les films plus

contemplatifs se concentrent sur l'expression émotionnelle et la précision du dialogue. Un autre facteur important à considérer est l'environnement de tournage, qu'il soit en extérieur ou en intérieur. Cela peut influencer plusieurs choses : la portée de voix, la position des panneaux acoustiques, la manière de percher etc.

Perspectives

Il y a de nombreux aspects de la post-synchro qui ne font pas l'objet de ce mémoire. Du côté des relations humaines, il serait pertinent de discuter avec les comédien·ne·s concernant leur perception de la post-synchro. Au cours de nos recherches pour ce mémoire, nous avons souvent entendu que certain·es·s comédien·ne·s français·e·s n'apprécient pas la post-synchro, car iels considèrent cette étape comme le résultat d'une performance insatisfaisante ou un problème technique lors du tournage. Leur donner la parole nous permettrait de mieux comprendre leurs besoins et leurs appréhensions vis-à-vis de cette séance en auditorium.

Il peut également être intéressant de tester différents types de microphones pour la post-synchronisation. Les perchman et perchwoman n'ayant pas toujours un casque, cela pourrait être pertinent de tester les changements de timbre et de couleur par rapport à l'angle et la distance de prise de son. Cela permettrait de mieux anticiper les positions de la perche en séance de post-synchro.

Enfin, il y a d'autres techniques de post-synchro en cours d'application. Par exemple, pour les séries, une solution intéressante a récemment été mise en place par la société Mont-Mars à Marseille³¹ : un camion mobile d'enregistrement permettant de faire la post-synchro et à anticiper la post-production dès le tournage. Ce cas pourrait être pratique pour une série qui tourne plusieurs épisodes ou saisons dans les mêmes lieux. Il pourrait être intéressant d'étendre l'étude de l'intégration de la post-synchro sur l'ensemble des productions audiovisuelles, surtout aux cas avec peu de budget mais des exigences élevées.

³¹ <https://www.mont-mars.com/>

BIBLIOGRAPHIE

- Avarese, J. (2017). *Sound Design: The Art and Craft of Audio Post Production for The Moving Image*. Oxford: Focal Press.
- Blanchet, A. & Gotman, A. (2010). *L'enquête et ses méthodes : L'entretien* (2ème édition). Armand Colin.
- Chion, M. (1982). *La Voix au cinéma*. Éditions de l'Étoile. Paris, France
- Chion M. (1985). *Le Son au cinéma*. Éditions de l'Étoile. Paris, France.
- Chion, M. (1994). *Audio-vision : Sound on Screen*. Columbia University Press.
- Deshays, D. (2010). *Entendre le cinéma*. Paris, France : Klincksieck.
- Dutrieux, L. (2021, 12 décembre). *Réenregistrer dans la couleur et le raccord : la postsynchronisation*.
- Kerins, M. (2010). *Beyond Dolby (Stereo) Cinema in the Digital Sound Age*. Indianapolis: Indiana University Press, p257, p297
- Lerouge, C. (1996). *Sur 100 années, le cinéma sonore*. Dujarric.
- Paillé, P. & Mucchielli, A. (2021). Chapitre 12 : L'analyse thématique. Dans P. Paillé & A.Mucchielli (dirs.), *L'analyse qualitative en sciences humaines et sociales* (p.269-358). Armand Colin.
- Pommier, C. (1988). *Doublage et Postsynchronisation*. Dujarric.
- Purcell, J. (2013). *Dialogue Editing for Motion Pictures : A Guide to the Invisible Art*. Taylor & Francis.
- Shukla, G. (2022). *La bande-son dans le cinéma italien des années 1960-70 : la recherche d'une réalité subjective* (Mémoire de master, ENS Louis Lumière). Paris, France.

ANNEXES

I- FORMULAIRE DE CONSENTEMENT

N° d'identification du participant :

Titre du projet

RECHERCHE QUALITATIVE auprès de :

Introduction :

Ce mémoire cherche à étudier le processus de l'intégration de post-synchronisations dans une scène. Il cherche à faire un état des lieux des techniques anciennes et actuelles utilisées à ce but. Nous allons à la rencontre de professionnels du métier ayant un rapport particulier avec les post-synchros pour mieux comprendre le travail effectué aujourd'hui ainsi que les contraintes actuelles pour nourrir le mémoire ainsi que la partie pratique du mémoire.

Les objectifs de ce projet sont d'explorer :

Le workflow de la post-synchro

La question de la vraisemblance

La finalité est d'obtenir des témoignages permettant de représenter le travail de post-synchro actuel.

Réalisation de l'entretien :

Cet entretien sera réalisé par Sreesti Sheikh suivant vos disponibilités.

Il durera de 30 à 90 minutes et sera enregistré de façon anonyme si souhaité.

Qu'est ce qui se passe si je participe ?

Vous participerez à un entretien individuel où l'on vous posera des questions concernant *votre expérience de :*

Vous avez la possibilité de quitter l'étude à n'importe quel moment sans fournir d'explication.

Comment sera traitée l'information recueillie ?

Les enregistrements seront retranscrits mot à mot de façon anonyme et confidentielle si souhaité.

Une fois transcrits, les enregistrements seront détruits. Les transcriptions seront gardées de façon sécurisée.

L'analyse des données sera réalisée par Sreesti Sheikh.

Les résultats seront utilisés dans le cadre de la thèse L'intégration de post-synchros, mémoire de 3^{ème} année de l'ENS Louis Lumière et peuvent éventuellement être publiés.

Merci de noter vos initiales dans chaque case :

1. Je confirme avoir lu et compris l'information ci-dessus et que j'ai eu la possibilité de poser des questions.
2. Je comprends que la participation est entièrement basée sur le volontariat et que je suis libre de changer d'avis à n'importe quel moment. Je comprends que ma participation est totalement volontaire et que je suis libre de sortir de l'étude à tout moment, sans avoir à fournir de raison.
3. Je donne mon consentement à l'enregistrement et à la transcription mot à mot de cet entretien.
4. Je donne mon consentement à l'utilisation éventuelle de certaines citations de l'entretien dans une thèse ou dans une publication.
5. Je suis d'accord pour participer à l'étude.
6. Je souhaite rester anonyme.

Signature (participant)_____

Signature
(investigateur)_____

Date_____

Date_____28/03/2023_____

Nom_____

Nom_____Sreesti Sheikh_____

II- QUESTIONNAIRE PPM

Votre expérience dans le domaine du son :

PARTIE 1

SEQUENCE X

- Pour chaque extrait, évaluez la vraisemblance de la voix de 1-5. On dira qu'une voix post synchronisée est vraisemblable s'il est difficile de la différencier du son direct.
- Indiquez si d'après vous, c'était que du son direct.

Extrait	Vraisemblance	Son direct		
	<i>1 : peu vraisemblable</i> <i>5 : vraisemblable</i>	<i>Encerclez votre choix</i>		
A		Oui	Non	Je ne sais pas
B		Oui	Non	Je ne sais pas

- Encerclez l'extrait que vous avez préféré : A B Je ne sais pas
- Merci d'expliquer pourquoi :

PARTIE 2

SEQUENCE X

Vous allez voir 3 mixages du même extrait. Notez indépendamment chaque extrait selon les critères du tableau.

Mixage	Vraisemblance	Acoustique de la voix	Vivacité de la voix
	<i>1 : peu vraisemblable</i> <i>5 : vraisemblable</i>	<i>1 : les voix ressortent</i> <i>5 : voix bien intégrées dans</i> <i>l'espace</i>	<i>1 : voix statiques</i> <i>5 : voix vivantes</i>
Mix 1			
Mix 2			
Mix 3			

III- Boom-It : entretien avec un des développeurs, Antoine Martin

Pour te mettre un peu dans le bain je vais te demander juste de te présenter, et tu peux aussi présenter Hal.

Je m'appelle Antoine Martin, j'ai fait Louis Lumière son et je suis sorti en 2016 et je me suis dirigé tout de suite vers du montage son. J'ai fait plein pas mal de courts métrages et de longs métrages en assistant et en parallèle de ça grâce à un stage que j'avais fait au cours de mes études chez UV qui est une boîte qui fait du développement de logiciels et en particulier qui développe des émulations de synthétiseurs vintage j'ai appris à développer et donc en parallèle de faire du montage son j'ai commencé à développer des plugins, dont Walker, qui est un plugin de bruit de pas que tu joues avec ton clavier midi. J'ai fait un truc s'appelle Meteor pour faire des impacts et du coup j'ai fait 70% de montage son et 30% de développement pendant plusieurs années. En 2020 j'ai intégré HAL, donc c'est une société de post-production avec des monteurs sons comme Nicolas Becker, des mixeurs comme Cyril Holtz et j'intégrais HAL pour faire particulièrement du développement chez eux. Au début j'ai commencé à mi-temps et maintenant je suis à temps plein et je fais plus que du développement de plugins, de logiciels et aussi de la création de bande son, donc je fais plus du tout de film en tant que monteur son depuis quelques mois.

Pour l'instant je suis en pause sur les films parce que j'ai envie de me consacrer à fond à cette activité là pour voir où c'est que je peux l'amener et j'aimerais bien j'imagine un jour reprendre les films.

Comment l'idée est venue de développer un logiciel comme Boom-IT, et à quel problème ça répond ?

L'idée derrière cet aspect développement chez Hal c'est que c'est les utilisateurs et les utilisatrices qui donnent les idées, c'est ça qui est vraiment la philosophie de la marque. Boom-it c'était une idée de Cyril notamment et de Lucien Richardson qui est associé de Polyson aussi et qui fait partie de la famille HAL depuis quelques années et donc ils voulaient un outil pour intégrer facilement les post-synchros. L'idée est venue quand Cyril bossait sur *Fireheart* (*Vaillante* en français), c'est un film d'animation avec une petite fille qui veut devenir pompier

qui est sortie il y a un an ou 2. Tous les doublages du film ont été enregistrés avec un micro devant le comédien, la comédienne et toujours à la même position, sauf que dans le film les personnages se déplacent dans l'image se retournent sur eux-mêmes, s'éloigne etc. Le son a toujours été enregistré comme ça, avec un 87 et on s'est demandé comment transformer cette prise de son assez facilement en quelque chose qui matcherait les mouvements qu'on a à l'image. C'est venu de cette idée-là et après on découlait d'autres idées comme pouvoir simuler plusieurs micros de tournage différents, pouvoir simuler tous les mouvements de perche possibles en plus des mouvements de personnages, comme si tu pouvais simuler un perchman par exemple qui bouge un petit peu un perchman qui détimbre etc., mais aussi simuler une personne qui se retourne par rapport à la perche. Les mouvements de personnages et les mouvements de perche c'est forcément lié parce qu'une perche qui se recule ou un personnage qui se recule par rapport à la perche en fait c'est la même chose. C'est juste ça le but de Boom-It quoi c'est recréer du mouvement dans une prise de son qui a été enregistrée de manière statique.

Une fois que vous saviez ça comment vous avez élaboré le logiciel ?

On s'est demandé ce qui se passait physiquement quand un personnage se retourne au niveau du son donc il y a la couleur donc à priori de l'EQ, il y a aussi la dynamique quand tu te recules tu tasses les transitoires. Les bruits de bouche les choses comme ça c'est typiquement de l'effet de proximité quand t'es loin t'en as moins. Il y a de l'acoustique aussi typiquement quand tu te retournes t'as moins de champ direct et t'as plus de diffus, et en fait nous on a pris le parti pris de faire que de la couleur et pas du tout d'acoustique parce qu'en fait l'acoustique pose plein d'autres problèmes : il faut avoir un moteur de convolution qui est intégré ou tu viendrais choisir ton espace ; c'était beaucoup trop de développement vu la taille de la structure et de nos moyens. Si les gens ont envie de donner de l'acoustique ils mettront leur réverbération derrière. On n'a même pas fait non plus de compression, ça c'est pour Boom-It 2 à mon avis. Boom-It à l'intérieur c'est juste une dynamique et les courbes d'EQ variant en fonction des mouvements des personnages et de la perche.

Techniquement comment vous avez fait ?

On a fait des IR, donc on est allé à l'IRCAM dans la chambre anéchoïque et on a loué un *torso simulator*. C'est un buste avec un haut-parleur à la place de la bouche donc c'est un outil de

mesure, ça a la taille vraiment d'un buste d'être humain. On s'en est servi entre autres pour diffuser des *sweeps* avec le petit haut-parleur à la place de la bouche. La tête elle est pivotable, donc on avait placé une perche par rapport à ce buste et on a fait tourner la tête, tourner le buste et pencher la tête etc. dans toutes les positions par rapport à un micro de mesure. On avait habillé notre buste pour éviter des réflexions sur son torse. On lui a enfilé un pull et un petit bonnet et ça simulait des réflexions plutôt naturelles. Notamment quand tu penches la tête vers le bas ta voix qui devient un peu plus grave etc. on a pu simuler tous ces mouvements. On a fait des mesures à 360° et avec les positions de la tête genre à 30° vers le bas et 50° vers le haut.

Je vais être un peu plus précis. Vous avez effectué des mesures d'angle, mais comment se passe la transition entre ce qui se passe, par exemple, à 30° et à 0° ?

Oui, il y a aussi des mesures où c'est un changement progressif. Pour ne pas dire de bêtises, nous avons simplement effectué des mesures à moins 30°, 0° et plus 50°. Ensuite, ce que nous obtenons, c'est que les valeurs, après les avoir converties en courbes EQ dans notre système de match que nous avons développé en interne, créent une sorte de correspondance entre les réponses. On obtient donc une courbe d'EQ, et on passe ensuite de manière fluide d'une courbe à l'autre. Donc, en fait, lorsque vous avez une courbe EQ pour une position et une autre pour une autre position, vous pouvez voir, par exemple, qu'il y a une bosse à 0° qui n'est pas présente à moins 30°. Donc, quand on incline la tête vers le haut, on voit cette bosse apparaître, par exemple. Ça se fait de manière linéaire, exactement, avec chaque EQ représentant une trentaine de points en termes de précision. Entre deux positions, nous examinons ce que devient chaque point dans la position suivante, ce qui nous permet d'interpoler de manière totalement fluide.

Au début de Boom-It on avait une visualisation des courbes et on voyait les qui simulaient on va les détimbrages. On a aussi fait des IR pour simuler les EQ des micros, donc on avait notre enceinte coaxiale qui diffusait des *sweeps* et on avait notre micro, on a pris plein de micros de tournage et on les a testés. Parce que par exemple quand on simule un micro semi canon ou un micro cardio c'est pas du tout pareil les effets de détimbrage et donc dans Boom-It tu peux sélectionner si tu veux simuler un micro cardio ou un semi canon ou un hyper cardio et le tilter.

Que ce soit la personne qui bouge ou que ce soit le micro qui bouge, c'est pareil ? Sinon quelle est la différence entre les deux ?

Alors en fait, le "tilt" du micro est assez spécifique. En tournage, le micro n'est jamais parfaitement aligné avec la bouche de la personne qui parle. Le micro, tu le sais toi-même, est souvent positionné au-dessus. Donc, par rapport à l'axe de la bouche, il n'est pas exactement en face, il est toujours un peu comme ça (il fait un geste). C'est pour cette raison que nous avons ajouté cette fonctionnalité. Mais en soi, ce n'est pas tout à fait pareil. Comment dire... Attends, je réfléchis à comment te répondre de manière claire. En gros, si nous avons enregistré seulement comme ça (il fait un geste), et que tu veux recréer un tilt en baissant la tête, il ne se passera pas la même chose que si tu es comme ça (il fait un autre geste) et que tu utilises le micro de cette manière. En fait, les effets lorsque tu baisses la tête créent vraiment quelque chose avec le torse. Donc, nous avons dû enregistrer tous les mouvements de tête et tous les mouvements du micro, car cela ne crée pas exactement les mêmes effets. Par contre, en ce qui concerne les distances, c'est vraiment similaire. Par exemple, lorsque le percheur est là et que ta perche est là et que tu fais ça (il fait un geste), ce qui se passe là-bas, c'est exactement la même chose. Mais en ce qui concerne les effets de tilt, tu étais obligé de faire les deux.

Et par rapport à l'éloignement, combien de mesures avez-vous effectuées ?

Nous en avons fait trois. Nous avons pris en compte différentes distances. En ce qui concerne les micros de tournage, nous n'avons pas simulé différentes distances. Il s'agit simplement de... Attends que je ne dise pas de bêtises... Je me souviens que nous avons réalisé des centaines de mesures il y a deux ans. Nous avons effectué trois distances, je crois, allant de 50 cm jusqu'à plusieurs mètres, environ 3 ou 4 m. Je ne me souviens plus si nous avons utilisé tous les micros pour cela, car certains étaient des micros de mesure, comme les micros infrarouges (IR) pour le torse. Ces micros de mesure étaient utilisés pour obtenir les effets des réflexions et des variations de timbre. Ensuite, nous avons utilisé les micros de tournage, un par un, en les plaçant devant un haut-parleur et en enregistrant les différents champs sonores. Oui, effectivement, nous avons utilisé plusieurs distances à ce moment-là. Nous avons également utilisé des micros de studio, car dans *Boom-It*, il y a une fonctionnalité permettant d'atténuer la coloration du micro utilisé lors de l'enregistrement. Par exemple, dans le film d'animation, nous avons utilisé un micro 87 avec un effet de proximité, et nous avons pensé qu'il serait intéressant de rendre le son d'origine neutre, puis de lui donner la coloration du micro utilisé lors du tournage. C'est ainsi que cela fonctionne dans *Boom-It*. Il y a plusieurs modules, dont un appelé *Studio Mic* où tu choisis le micro utilisé, puis un autre module appelé *Set Mic* où tu

choisis le micro que tu veux simuler. Donc, quand tu choisis un *Studio Mic*, ça ne simule pas le son qu'il produit, mais les caractéristiques physiques. Par exemple, pour le micro 87, nous avons sa courbe de réponse, et nous pouvons voir s'il y a une augmentation des basses à 80 Hz ou quelque chose du genre. Dans Boom-It, nous atténuons cette caractéristique comme si nous voulions rendre la prise de son neutre, puis le deuxième module donne la coloration correspondante. Je comprends que ce n'est pas très clair pour beaucoup de gens, donc nous pensons à améliorer l'interface de Boom-It pour rendre cela plus compréhensible. En fait, je suis en train d'y réfléchir. Pour l'instant, ce n'est pas du tout abouti, mais laisse-moi te montrer. Nous réfléchissons à une nouvelle interface pour Boom-It, qui sera encore... Enfin, pour l'instant, on dirait que ça n'a pas beaucoup changé, mais cela a déjà un peu changé par rapport à la version que tu connais. Par exemple, nous avons réduit... Attends, je vais te montrer l'ancienne interface. Voilà, avant, c'était comme ça (il montre une ancienne interface), et donc là, nous pensons qu'il y a beaucoup d'espace, par exemple ici, énormément d'espace, alors qu'il y a juste un seul bouton pour le "liveliness".

Je parlais avec un mixeur qui a utilisé Boom-It et il trouve ça très bien, mais c'est vrai qu'il n'avait pas compris tout de suite ce que faisait le module avec les différents *Set Mic* et *Studio Mic*. Surtout le fait que par défaut le *in-out* ça soit déjà configuré, parce que la couleur changeait alors qu'il avait rien fait comme réglage C'est possible de penser que ça prend le son d'un micro pour le transformer en un autre.

Oui, ce sont des trucs comme ça que les gens pensent. Mais tu fais bien de me le dire, je prends les retours en tout cas et tiens d'ailleurs dans la nouvelle interface (il montre la nouvelle interface), on va changer ça. J'ai aussi éclairci un peu le fond pour qu'on voit mieux la perche. Il y a un petit bouton info que j'ai mis là, parce que je trouve ça bien dans les plugins quand tu peux cliquer dessus et avoir une sorte de manuel qui est directement intégré dans l'interface. Il permet de savoir ce que fait chaque module. Mais je suis d'accord avec toi, la partie micro elle n'est pas claire. Même moi-même en essayant d'expliquer je me rends compte que je suis obligé de parler de ça pendant 3 min donc ça prouve que c'est pas encore super clair. C'est un peu une erreur de...on va dire une erreur de jeunesse, c'est à dire c'est notre premier plugin donc voilà on apprend de nos erreurs. Je pense d'ailleurs que l'intérêt de Boom-It c'est pas ça c'est plus le bonhomme, les mouvements et le liveliness, on mettra ces module en avant dans la nouvelle interface.

D'ailleurs comment fonctionne le liveliness ?

Le liveliness c'est les détimbrages un peu chaotiques qui sont appliquées sur tous les paramètres, que ce soit le tilt du micro ou la position du personnage. Donc ça va juste osciller entre la position que tu as choisi pour le personnage, ça va osciller quelques degrés et plus tu augmentes le liveliness et plus oscillations elles sont audibles sachant qu'elles sont dans tous les cas aléatoires.

Par rapport à l'utilisation de *Boom-It* dans le workflow de mixage, tu le comparais plutôt à des EQ ?

Oui en fait à l'intérieur c'est une EQ. Je le compare à une EQ sauf que pour obtenir ces résultats aujourd'hui enfin sans Boom-It tu fais des high shelf ou des choses on va dire assez simples, et tu vas faire à l'oreille ce qui est très bien aussi de fonctionner à l'oreille, mais sauf que là avec *Boom-It* c'est un vrai rapport visuel. L'idée c'est de matcher les mouvements qui se passent à l'image. Tu prends ton personnage et puis tu fais intuitivement le même mouvement, sachant que c'est un personnage 3D donc tu peux cliquer dessus, le déplacer dans tous les sens, il peut baisser la tête la monter, tourner etc. Ça c'est assez c'est un petit challenge technique, c'est qu'on est le premier plugin à utiliser un bonhomme en 3D, d'habitude les plugins c'est toujours la 2D. On a utilisé le moteur graphique open gel et du coup on a le premier plugin avec un bonhomme en 3D donc c'est cool et c'est assez intuitif à utiliser. Tu peux faire sans Boom-It avec des EQ manuelles mais c'est moins précis, plus long et moins drôle.

Aussi en termes de rendu et de qualité, c'est juste une question de...justement de vraisemblance. On voulait justement le rendu le plus, même pas vrai semblant, mais le plus proche de la réalité possible. Nos EQ ne sont pas juste des shelves et des choses comme ça, c'est des EQ avec plein de points et des bosses dans tous les sens et qui représentent réellement ce qu'on a mesuré en chambre anéchoïque. Après il faut savoir que dans la réalité en fait les effets d'éloignement ce qui est le plus audible je pense c'est l'acoustique. La différence majeure c'est qu'en fait on entend quasiment plus de champ direct quand je me tourne (il se tourne). Du coup les effets juste sur la couleur sont assez subtils, c'est pour ça qu'on a rajouté ce petit potard-là qui s'appelle *scale*, qui augmente ou diminue les effets de l'EQ.

C'est quoi la différence entre « gain » et « gaincomp » ?

Ça j'aurais bien aimé qu'on mette le titre en entier effectivement parce que ça permettrait de mieux comprendre, c'est *gain compensation*, et donc la compensation de gain. Notamment sur la distance ça va être pratique parce que quand tu recules le bonhomme, non seulement tu perds des aigus, mais tu perds aussi du niveau et donc une compensation ça permet par exemple un changement de couleur quand tu recules et pas de perdre du niveau.

Sans *gain compensation*, on a une vraie perte de niveau cohérente avec ce qu'on a mesuré en chambre anéchoïque, où la perte de niveau assez drastique donc tu entends quasiment 3 fois moins fort. Dans d'autres lieux plus réverbérés la perte ne serait pas aussi drastique donc c'est un bouton assez pratique.

Ce serait quoi le workflow idéal, part rapport à l'interaction entre Boom-It et une réverbe ou d'autres types de logiciels ?

Idéalement c'est bien d'avoir une réverbération juste derrière Boom-It, et quand tu augmentes la distance t'augmentes justement ton *Dry/Wet* de *reverb*. Dans le meilleur des mondes il aurait même fallu qu'on ait un petit moteur de convolution dans Boom-It, et qu'on sélectionne sa pièce et que la distance augmente les réflexions aussi, ça serait vraiment idéal. Mais ça aurait imposé de coder une réverbe et du coup c'était un an de développement en plus.

L'intérêt de mesurer différentes distances en chambre anecdotique c'est juste pour avoir l'influence de et de la distance sur la couleur du son mais pas sur les réflexions et donc ça c'est assez intéressant, c'est quelque chose qu'on mesure rarement en fait parce que justement quand tu as une prise de distance tu as plein de paramètres qui jouent autre que la couleur. Boom-It c'est vraiment un outil de couleur donc pas inintéressant, mais il faut ajouter d'autres outils pour avoir quelque chose de cohérent dans l'espace.

Vous avez eu des retours d'expérience ?

Oui mais pas tant que ça, pour être 100% transparent on n'a pas vendu des milliers de licences, on en a vendu environ 150. Parmi ces 150 personnes qui l'ont acheté, on a eu quelques retours. Beaucoup de retours nous disent que c'est super, et d'autres qui parlent de l'interface des différents *mics* qui n'est pas très clair. L'idée serait de le mettre à jour-là dans l'année pour justement prendre en compte ces retours, mais en tout cas vraiment on a des très bons retours sur la partie centrale, à savoir la partie *head tilt*, angle et distance. On obtient des résultats en deux clics, c'est ça qui est intéressant aussi, c'est le côté simplicité, intuitivité et il faut vraiment pas négliger le côté ludique aussi. Les plugins c'est un peu des jouets pour nous, donc il ne faut pas que ça soit chiant à utiliser.

Tu t'es inspiré de certains logiciels pour l'interface graphique ?

L'interface c'est pas moi qui l'ai désignée, je l'ai pensé au début mais c'est un graphiste qu'il a fait. On voulait le bonhomme soit non genré au début, et maintenant il a quand même une silhouette un peu masculine mais on se disait, ce bonhomme il faut qu'il puisse symboliser tout type de personnage. Il y a peut-être une update à faire parce que là il est encore un peu carré... Le graphiste on l'avait eu pour 3 jours donc on a pas pu peut-être aller totalement au bout de ce qu'on voulait faire graphiquement. En termes d'influence je dirais que FabFilter c'est un peu de génies, je trouve leurs plugins super simples et vachement clairs, c'est à dire que quand tu l'ouvres tu sais comment tu vas t'en servir. J'ai envie que les plugins qu'on fasse avec HAL soit comme ; tu l'ouvres et très rapidement tu fais ce que t'as envie de faire sans avoir à regarder le manuel.

Y a-t-il d'autres problématiques ou d'autres problèmes qui ont été mis en lumière lors du développement de ce logiciel ?

Oui, en ce qui concerne la dynamique, par exemple, il y a une atténuation de proximité ici qui permet de réduire l'effet de proximité que l'on peut avoir sur une prise de studio. Donc, si tu as enregistré ton instrument ici, tu auras forcément cet effet de proximité, et cette fonction permet de l'atténuer. Mais ce n'est pas dynamique, c'est simplement une atténuation des graves. Donc, il faudrait ajouter de la dynamique pour que ce soit vraiment naturel et intéressant. De même, lorsqu'on s'éloigne, il n'y a pas de compression des transitoires, ce qui devrait être le cas car lorsque tu t'éloignes, ta voix se comprime légèrement. Donc, ça n'a pas été intégré, mais il faudrait le faire.

Et, pour moi, je ne sais pas pourquoi, mais dans le processus de création, il y avait des variations de niveau. C'est inévitable car le "liveliness" joue sur tous les paramètres que tu vois ici, que ce soit l'inclinaison du micro, la distance, l'orientation de la tête, et tous ces paramètres varient. À moins que tu aies mis la compensation au maximum, il y aura forcément des variations de niveau, ce qui entraîne des variations de couleur.

Le problème d'avoir un compresseur à l'intérieur de notre moteur, c'est que cela pose des problèmes de niveau d'entrée et donc de seuil, parce qu'un compresseur doit être réglé et le seuil doit être optimisé. Mais du coup il faudrait analyser le niveau d'entrée pour déterminer le seuil si l'on veut que cela fonctionne de manière automatique. Sinon il faudrait mettre un bouton de seuil accessible à l'utilisateur. Je ne sais pas, FabFilter le fait assez bien : dans la version 3, il y a une variation de dynamique où le seuil s'ajuste automatiquement en fonction du niveau d'entrée détecté. Je dirais que ce serait l'idéal de fonctionner de cette manière, mais c'est compliqué.

Ce qu'il faut c'est repérer le niveau d'entrée et mettre un seuil qui corresponde à un niveau d'action du compresseur un peu léger. Il faudrait que ce seuil s'adapte au niveau d'entrée, plus le niveau d'entrée est fort plus le seuil est haut de sorte à ne pas trop compresser, ce serait pas naturel. Je pense que laisser cette option là l'utilisateur s'est rajouter encore des contrôles et peut être potentiellement perdre l'utilisateur. Il faut que ce soit caché pour qu'on ne soit pas vraiment conscient qu'il y a un compresseur en regardant l'interface.

Des choses à ajouter concernant l'avenir du logiciel ?

Oui, pour l'avenir, clairement, il y aura une mise à jour graphique. Quand il est sorti, j'avais l'impression qu'on aurait pu aller plus loin, mais faute de temps et de moyens, on a sorti cette version graphique. Maintenant, avec les retours que nous avons, il me semble important de les intégrer, de faire évoluer le logiciel. En termes de mises à jour techniques, nous allons d'abord résoudre les petits bugs que nous avons rencontrés, puis assurer une compatibilité native avec Apple Silicon. Voilà, il n'y a pas de grandes fonctionnalités que j'envisage d'ajouter immédiatement. Potentiellement, un compresseur, voilà. Mais par exemple, pas d'acoustique, du moins pas tout de suite. Même si c'est une fonctionnalité intéressante, c'est trop compliqué.

Je pense que cela rendrait le logiciel trop complexe à gérer, ajouter un moteur de convolution à l'intérieur de Boom-It, à méditer, voilà.

Tu avais de l'expérience en programmation avant de venir ici, non ?

Ma première expérience était en stage chez UVI. Mon projet de mémoire était Walker, donc un truc de bruit de pas que l'on joue au clavier MIDI. Je leur ai présenté mon patch Max à l'époque, et je leur ai montré, tu vois, on peut jouer des bruits de pas, c'est rigolo. Ils m'ont dit : "Ah, ce serait génial si tu le maquettais toi-même, que tu apprennes à coder, que tu nous fasses une preuve de concept." Donc c'est ce que j'ai fait, j'ai appris à coder en 3 semaines, tu vois. Mais bon, je dis ça, j'ai appris un seul langage, ce qui est tout à fait faisable en 3 semaines si tu t'y mets à fond.

Est-ce que tu vois d'autres logiciels arriver ou qui manqueraient pour l'intégration des post-synchros ?

Je faisais surtout du montage avant de me consacrer au développement. Donc, en fait, j'ai fait très peu de post-synchro, je ne suis pas forcément le mieux placé pour te répondre. En tout cas, ce que je peux te dire, c'est vers quoi on se dirige en termes de technologie pour l'avenir lié au post-synchro. Il y a notamment toutes les technologies de « text-to-speech » qui sont en train d'exploser et qui deviennent de plus en plus réalistes. Elles vont vraiment changer le monde de la post-synchro, notamment avec des solutions comme Voice Designer développé par Soundly. C'est un outil de conversion texte en parole. Par exemple, tu peux taper une annonce d'aéroport, du genre "embarquement porte B11", et tu sélectionnes un comédien ou une comédienne virtuelle qui dira le texte de manière hyper réaliste dans la langue de ton choix. Tu peux également choisir des effets, comme un effet de haut-parleur d'aéroport avec une acoustique. Donc, tu peux générer ce que tu aurais fait auparavant avec un comédien ou une comédienne en post-synchro, et ça changera considérablement la post-synchro. Il y a aussi toutes les techniques de morphing vocal, qui permettent de transformer la voix d'une personne en celle d'une autre, c'est en plein développement. Par exemple, si tu veux juste refaire un mot, imaginons que tu aies Gérard Depardieu dans ton film et que tu veuilles juste refaire un mot, est-ce que tu vas payer une demi-journée de post-synchro à Gérard Depardieu pour un seul mot, ou est-ce que tu génères ce mot manquant à partir d'une IA qui a été entraînée sur sa voix, et tu économises de l'argent ? C'est possible aujourd'hui.

Clairement, ce sont des technologies dont on va parler de plus en plus. En ce moment, c'est extrêmement à la mode, tu vois, il y a une effervescence, parce que comme tout ce qui est à la mode, tout le monde en parle. Ça va sûrement se stabiliser un peu, mais ça va vraiment modifier nos manières de travailler, notamment pour tout ce qui concerne la post-synchro d'ambiance, la post-synchro des figurants, tu vois, le couple qui parle en arrière-plan dans le restaurant, les annonces dans les aéroports, etc. Toutes ces voix-là que nous faisons aujourd'hui avec des comédiens, je pense qu'on ne les fera plus autant. Enfin, je ne sais pas si c'est souhaitable, ce n'est pas la question. Je ne sais pas si c'est une bonne chose, mais en tout cas, c'est ce qui se profile.

Merci beaucoup.

IV- Dialogue Match : entretien avec les développeurs de chez iZotope

Mike Rozett (Principal Product Manager), Nick LaPenn (software developer, head of development, Dialogue Match)

And basically, I'm working on ADR matching in my thesis, so I'm just going through how ADR has been recorded and mixed through the years and then what's happening now in terms of what the professionals are using and what's being developed. So how did you get the idea for Dialogue Match? Can you go through the timeline of it?

Mike

Yeah, maybe Nick, I'll kick off with the idea, and then hand off to you for how we did it. But for us, we were always talking to people in the industry, and you know, we have a beta program. We have friends over the years who call us up and say, "Hey, there's a problem. Can you help us with that?" And one of the things that came up was a problem that, I'm sure from your research, you're very familiar with now, but it was one of those situations where people, particularly dialogue editors, were being handed dialogue tracks that were recorded on body mics or lavalier mics under clothing. And there was also the boom mic, which obviously picks up a ton more room sound. And then if they didn't quite get what they needed from those two recordings, they would then go and have people do ADR in the studio. And you would get this classic situation where there were great performances on all three of those recordings, but no single recording had everything you needed.

And so, intercutting between the warmth and closeness of a lavalier with the large boomy room sound, which sometimes is fantastic, and other times it's a little bit out of control, and then taking a really clean, perfectly recorded read in a studio with ADR, but where it sounds maybe a little bit sterile and doesn't have the same characteristics of a live performance in a live space. How can you cut those things together and make them all seamlessly work, so you can pick the best of the lavalier, the boom, and the ADR, and make it a single dialogue track that feels like one cohesive performance? One performance that's believable.

So, people in the industry were saying to us, "Hey, you know you've done stuff with matching EQ. You've done stuff with reverberation, reverb, and you know how to capture and apply

ambiance. Is there something you could do with all of those things?" And Nick, I'll hand off to you to kind of talk about how we went after the solution.

Nick

So, what's interesting is that people have been asking us for something like this for a really long time. And if I remember correctly, the original use case may have been even broader. People really just wanted to be able to match any dialogue to any dialogue. They came to the team and said, "What can we do for this?" And we had some existing technologies and ideas about how to do it. But a lot of the process of developing this product involved narrowing down the use cases to the ones we could actually address because some use cases are trickier than others. Certain types of noise are easier or harder to remove. Removing reverb, for example, is especially tricky. If someone recorded something in a cave and they want to remove the reverb, that might be quite challenging.

On the other hand, we have shipped a bunch of other tools that match different types of things before. So, there was a lot of work and discussion around the workflow. How can we make a tool like this, which is doing so many things at once, easy to use and intuitive for our users? We had to take everything that Mike learned from all these customers and their requests and synthesize it into something that we could deliver and that would be genuinely useful.

Oay, and when you talk about dry and wet audio on the reverb module, can you do a dialogue match for a wetter audio to a drier audio for example? Is that possible?

Nick

That's one of the limitations that currently exists in it is that we can't go from wet to dry. It can work if the reverb, like the wet reverb signal, isn't too different from the original. Sometimes, using EQ and tools like EQ match can cover it up, for example, or having a shorter, louder reverb can help mask a quieter, longer tail. However, at the time, our reverb technology was still young, so we didn't have anything that could effectively remove longer reverberant signals from dialogue. Since then, our technology has matured significantly, so it's something we do consider revisiting.

OK. So I don't know how much information you can give me as to the machine learning and the algorithm behind it, but could you dive a bit into that please?

Nick

And there was a paper published about how most of this works that I can try to find you a link to if you want.

Yeah, sure.

Nick

So, like, you could actually read the paper. But yeah, what we did is we really ran a lot of dialogue. Basically, we created our machine learning algorithm that tries to listen to dialogue with reverb and select the closest matching reverb preset. So, what we did is we essentially preprocessed a lot of dry dialogue with reverb. Then, we had some very talented machine learning researchers who spent a lot of time tweaking the architecture that we played that audio through. The algorithm tries to match it to the closest preset and is given a score on how well it performed. Through repetition, it aims to improve that score.

It's not super complicated, but yeah, I can try to find the paper for you because it might provide more detailed information. We're not doing anything completely groundbreaking. We are using existing machine learning architectures.

OK, so you're not generating an IR, it's matching to pre-existing presets to achieve the desired result, right? And once the machine outputs the result, do the researchers need to manually score the output on a scale of 1 to 10, or is there an automatic way to determine it?

Nick:

There are various ways to automate it, depending on the architecture. One example is if we have knowledge of the preset or settings that were used for processing, we can identify the important parameters and assess how closely the output matches them. For instance, if the original preset had a one-second tail and the matched result has an 0.8-second tail, it would be

considered 80% accurate. That gives us a good indication of the similarity. I don't recall the specific metrics we used for this off the top of my head. It was implemented several years ago.

I had a question about the EQ match and the ambience match modules that are on there; are they exactly the same as what we have in RX-10?

Nick

Yeah, so the ambience match module is the same as the static ambience match module we had at the time. That module is really nice as it can listen to a wet signal and extract just the ambience. One somewhat unique aspect is that you don't need ambience preservation to extract ambience with that module, which is pretty cool.

As for the EQ match, we tried various types. We have different tools that can perform exact EQ matching or match a general tonal curve, which we evaluated extensively. However, we ended up using the same EQ match technology as RX. If I recall correctly, it's not set up to be super surgical. Instead, it aims to adjust the overall tone of the signal rather than striving for an exact match between two takes. This approach is more suitable since users often work with short takes but apply the adjustments to many different ones.

Do you have any advice as to what the ideal audio could be to capture if, say, we want the three modules to be working together?

Yeah, I mean, I would say the best advice would probably be to capture multiple versions of both the source and the target. Obviously, you'll want to capture something that includes the ambience for the EQ. The more you can learn from, the better, as it ensures that what you're capturing represents a more average EQ curve of the clip. It's possible that if you capture a bit where the actor was making a funny voice, for example, that they wouldn't normally make, it won't match up.

Another thing to consider is that the reverb matching is based on machine learning. We taught it, but we can't look inside the box to know exactly what it learned, right? So if you run the reverb match on different clips, you'll get different results, and usually, they're pretty close. However, there will be some user preferences, and our users' ears will be really good at identifying which clip they think is the best. We also provide controls for the user to tweak settings because it's really difficult to achieve an exact match.

Right when you say multiple inputs, do you mean trial and error or is there a way to give it multiple clips?

Nick

I would say more trial and error. I do think like our power users, they probably get to a point where they trust it and do it on multiple clips. I know a lot of them that work with the same actors and the same TV show and things like that over and over again. Once they've made a preset, the set dialogue, that sound is not really going to change a lot. So, once they have those presets, they can probably just trust them.

OK, so about the reverb module specifically, is it something that isotope alone developed? Or was it a partnership?

Nick

Mike, you can take this one.

Mike

If you want, I was just going to jump in here. So, iZotope acquired Exponential Audio four or five years ago now and we brought over the work, all the products, the algorithms and the research and so forth that Michael Carnes, who had created Exponential Audio, had done. He actually worked with us for about two years before, so we brought over his technology and we've been able to use that in a couple of different products. The reverb came from Exponential Audio. Exponential Audio was a company we acquired, and we incorporated their reverb into our lineup. Nick can provide more details on any modifications we made, but that was the initial source, one of his specific reverb engines.

Nick

Yeah, the main idea behind using the reverb module was more so that people are using that to match reverb already in their workflows. They were doing it manually with their ears though, so they would literally load the reverb on the track. They wanted to match and find a preset and tweak settings until they had it sounding close and we figured that the best way to probably help their workflow if they're already using that reverb is to use that reverb ourselves, and we

were lucky enough to have access to it due to the acquisition. But yeah, we did have to do some modification, but it was more around making it work and training it than around getting it into the product.

Right, since you have a reverbe module, are there stereo or multi-channel instances of use of Dialogue Match even though dialogue is mostly mono? Is that possible?

Nick

I honestly can't remember if we blocked the use for multi-channel. We might have because it really is only designed to be used on a mono clip and so... if you think about it, what do you do when you have a clip with eight different tracks, right? Should we be treating each track individually? Should we try we be trying to take the average? Maybe it's a surround sound clip and we should be ignoring the LFE. Basically, like we didn't really know what users would be doing with it for multi-channel, so we tried to stick the most common-use case, which was mono.

In terms of feedback, has anyone talked to you about multi-channel uses?

Mike

I haven't, and to be clear, I'm not saying there may not be a use for that. But I can't think of a single time when somebody asked us for that. I mean, we really were assuming that it was kind of a mono workflow and structure, so to speak. So yeah, I don't remember a single multi-channel request for Dialogue-Match.

Okay, what kind of request do you usually get in terms of tweaking the on the software for Dialogue Match?

Mike

Nick, I'm going to present one point, and then I'm sure you have some thoughts on it. We've heard people discussing early reflections quite a bit, right? It's probably one of the most challenging aspects to capture and reproduce. For instance, in a use case where you're recording inside a car on location or even on a stage with a car prop, the dialogue editor wants to capture the distinct sound of being inside a car, with the glass and the quietness, and the short, confined room sound with early reflections. It's much more difficult to transfer those characteristics in

such settings compared to locations with more reverberation, like a church where people are having conversations or in a cave, as Nick mentioned earlier. These larger and more dramatic environments with longer reverberation tails are easier to capture and reproduce. Regarding early reflections, it's something that people have been asking about, and without delving too much into future plans, I'll simply say that it's an area we're still exploring. It's challenging, but it's certainly an interesting aspect to investigate further.

Nick

Yeah, like, can I just interject? Yeah, I have a list. Early reflections is one aspect that frequently arises. We did attempt to address it. Many early reflections are perceived as EQ, so we tried to incorporate them into our EQ match. However, there are slightly delayed early reflections that prove to be quite challenging to capture. So that's probably the top concern. The second point is something you already mentioned, which is removing reverb from a wet signal. We excel at going from dry to wet, but we don't effectively support transitioning from a heavily reverberated signal to a dry one, and that's unfortunate. It's likely a significant use case, meaning people may have to choose the wet signal over the dry one in certain situations when using our product. The third aspect is dynamic ambience. The ambience match module we implemented focuses on static ambience and can simulate background fans, but it struggles with replicating other periodic sounds and similar elements. Lastly, we receive numerous requests for a separate tool that allows users to adjust pitch, timing, and timbre of voices. It's somewhat related to the main problem, but users often desire a comprehensive solution in one place. Speaking of that, it's currently available only in Pro Tools, and non-Pro Tools users express their desire for it. There are reasons behind this limitation, but it's something we consider, as we receive numerous requests for it.

OK, allow me to come back to the early reflections. So why exactly is it so hard to capture the early reflections part of it?

Nick

I mean, this is like early in our journey of investigating machine learning, but it has just been really hard for us to teach a machine to listen to and understand early reflections and select a preset based on them. It tends to do a really good job with late reflections, but for some reason, we weren't successful in achieving something that was spot on or even close with early reflections. I think part of it is because, for us, the difference between a one-second reverb tail

and a 1.1-second reverb tail doesn't necessarily make a huge difference, but with early reflections, even a 1/10 of a millisecond can completely change the EQ of a signal. So, I think we're also just a lot more sensitive to the effect that early reflections have on a signal.

Mike

I don't know. And Nick, correct me if I'm wrong. I don't know of anybody who's nailed back yet. It's still. It still feels like it's a problem out there that and I'm sure we're not the only ones thinking about it. That still could use work and does need work. There's definitely still research and development to be done around that.

Nick

Yeah, we don't know of a one-click solution yet for all these parameters.

Mike

And that's one of the good themes to cover while we're talking about this quickly, and I'm sure your research has shown you this as well, but when folks are working in these situations, when you have professionals doing sound for television and film, they will always tell you there's never enough time and too much work. So being able to move fast is really crucial. And as Nick was saying, you know, like, can we have a one-click solution? Sure, if you want to spend 20 to 30 minutes and use a particular piece or maybe a couple pieces of software, you can do some really great work. And if you're doing restoration work, you probably have the time to do that. But if you're putting together, you know, if you're working on a series of ten one-hour shows or a two-and-a-half-hour film, and you're trying to get the dialogue right, you have to move. I mean, we've seen these folks work. They move so fast using tools like, of course, Dialogue Match, but also RX and other impressive tools from competitors. They're using these tools in an incredibly fast, tight workflow. So, anything we can do, or anybody in the industry, that's like a one or two-click solution that gets you really close, and you just do a little bit of tweaking and in a matter of seconds move on, that's always what we're aiming for. Because speed is such an issue, and the clock is ticking, and they have deadlines, and we and many others want to be a part of helping them achieve great results as fast as possible. So that's always part of what we're looking at. It's not just about a 20-minute solution, but is there a two-second solution? Maybe we use machine learning and a combination of other technologies to get people there, but that's always on our minds in everything we're doing. And some solutions

can be done just like that, and others, each year we get better and better at delivering faster solutions, but some take more time than others.

Yeah, the post-production time is really short everywhere.

Mike

Particularly the audio departments. You know, there's that thing of, like, go to your favorite film, watch your favorite film or your favorite television show, and turn the audio off and see how great that experience is. Then you turn it back on, and you have this sort of immersive experience. Well, the audio department is getting a tiny fraction of the budget, but it's probably 50 percent or more of the actual experience. And so those folks are always saying to us, "Man, they keep cutting us back with less and less time." So, it is kind of a problem in the industry for true.

Nick

You have a feeling with streaming, it's probably even worse. Now the turnaround time is just an instant.

Mike

We've heard too from folks who working on podcasts that if you're, you know, if you're doing a one-hour podcast, once it's recorded, you're basically given an hour and a half to do all the work, and hand it over. People have a few hours sometimes to turn things in. That's how tight it is on that side. Obviously, TV and film productions have more time, but yeah, time is always the issue. It's kind of unfair, but we're doing, you know, research on that right now and it's really true with podcast folks that they have sometimes a few hours to get something ready to go.

Nick

It's like for news-type podcasts where they're racing against everyone else to get the content out about something that just happened. So that tends to be where it is, and it's all about building a tool that people can learn to trust. With very few clicks, they can build their presets, and they know what those presets are going to sound like every time they hit render. It lets them work that much quicker.

That sounds very tight. Going through the list of feedback you got, you also talked about static ambience and pitch?

Nick

So, instead of static ambience like a fan, think about dynamic ambience such as trees blowing. It's a sound that changes over time and is in the background. Anything recorded outdoors, people would like us to be able to match it. If we can't match it, they would at least like us to be able to remove it, which we can usually do.

And regarding the feedback you've received and the development of new plugins and software, have any new challenges emerged in recent years related to ADR or dialogue in general?

Nick:

Well, there's the issue that arose during COVID when people couldn't be on site with all the best equipment. Many individuals had to record using whatever they had available in their homes. During that time, we received numerous requests on how to clean up audio recorded on a phone and similar devices. However, I'm uncertain how long-lasting that particular challenge will be.

Mike

Yeah, I wonder the same thing. I mean, we were definitely there. They're in videos of folks from famous TV shows sitting in their cars with the windows up, using their iPhone to record ADR. Actors don't want to come into studios generally because they hate the ADR environment a lot, most of them. So, if you can do it on your phone, great. And like Nick said during COVID. We literally heard stories of people going into their, you know, bedroom closets or getting under a blanket and just trying to get an acoustically dead space. It wasn't 200 lines of dialogue, it might be some pickup lines. Then they were just emailing it off to the dialogue editor and being like "good luck, just get that synced up and do what you need to do so". But I agree with Nick, you can't get away with that anymore, that's going away to a large extent. Sorry, I don't mean to be a clock person, but we, Nick and I literally have a meeting right now for another thing, we're dealing with, so we're going to need to drop off. Sreesti did we get close to your list of questions? Do you have what you need?

I just have one more small question about the future of iZotope and what kind of stuff you're looking to develop. Also, what do you think about generating voices through AI and similar technologies?

Mike:

Good question. Nick, I'll take the product one, and maybe you can address the generative aspect. We're still putting a lot of effort into exploring how machine learning can assist us and solve various problems, making it a versatile tool that can quickly handle a wide range of issues. Specifically, in terms of RX, as I mentioned earlier, we're looking into expanding our repair and editing technology to cater to podcasters, post-production professionals, music creators, and content creators. We want to reach out to a broader audience and address their specific needs. This is an ongoing focus for us.

Nick:

Yeah, addressing the generative aspect is a bit trickier due to our non-disclosure agreements, but I can share that generative AI has already been part of our strategy. In RX-9, we introduced a tool called Spectral Recovery, which is capable of restoring frequencies that were missing in a band-limited signal. If you haven't tried it in the latest version of RX, I highly recommend giving it a go. Record your voice, remove everything above 4 kHz, and see the impressive results it can achieve. So, I see AI playing a significant role in audio restoration and generation moving forward.

OK. Thanks a lot for your time.

V- Chameleon : entretien avec Benjamin Graf, Accentize

What was the process behind the development of this product?

The story started when one of our first products came out, it was called *De-room* and it was for removing reverberation. When developing it we spent a lot of time on different types of reverbs and how to characterize them and how to remove them in the end, so we had to estimate the reverb content present in a recording.

This product was finished at some point and then we got the idea that since we already had the information of the reverb present in the recording, we could also try to extract an impulse response from this information so it could be used. Rather than being removed it could be applied to other recordings. So, then we did some research in this direction. At some point we realized it worked well and we started developing the product and getting in touch with people who work in ADR regarding the features that should be added. This is how it was built.

When you developed Chameleon did you know it would be mainly used for ADR? What was the objective?

First, we developed the technology and then thought about how we could build a product out of it. ADR was one of the main targets even though it's also used in music production in some cases, but now it's used 80 % in ADR. We knew that reverb matching was an issue that had not yet been resolved and we put two and two together when we developed this technology, it fit well.

What kind of feedback did you receive?

At first the feedback was about the estimation, the type of reverbs that were useful and they were amazed by it. There is Dialogue Match from iZotope, but it wasn't really satisfactory for the reverb part. They liked chameleon. Then the feedback shifted to how it integrates into the workflow. In our research it was about how to get the impulse responses, but now we needed to build automation features. There was a lot of back and forth with studios telling us what's needed for a faster workflow and automation.

What needed to be changed?

Mostly for ADR, sometimes music, we needed to make it available as inserts in mixing busses or in the audiosuite to match a few clips. What was really important was the ability to automate and switch between profiles, and that was something we hadn't thought about before people started actually using it.

Did that open up new problematics and give you any ideas about new features?

For now, Chameleon is focused on an overall convolution reverb but current feature requests have asked for an early reflection part because people want to estimate early reflection and have some control over that part. Regarding workflow there are still automation requests. We try to integrate the requests. Accentize has developed other products like De-room or Spectral Matching. Chameleon has the best feedback and the widest use because it is kind of unique in the market.

How does it work?

It's based on deep learning like all our products. There's a neuron network that is trained on lots of example recordings, dialogue in different rooms. We give this input and then show the network the desired output, which is the impulse response we're looking for and the network has to learn the mapping between the input and output. This is done repeatedly for thousands of inputs; it takes a few days before the network acquires the knowledge needed. We then try to feed it a new recording where it can extract impulse responses from. After the estimation it works like a convolution. You can also extract the IR in another reverb plug-in if you want to, like Altiverb. But what's also pretty cool about chameleon is that if you estimate a profile then the resulting profile is a lot smaller compared to audio impulse responses (a few kb instead of a few mb). You can save all the profiles inside the session in a very compressed form. This is also achieved by a neuron network. You can decompress it if you extract the IR but if you stay inside Chameleon it is really memory efficient. Right now, memory is not an issue but what's useful is that if you transfer the session to another machine, the impulse responses don't need to be transferred separately, they're already inside the project. This depends on the neuron network which does not output the audio waveform but a compressed form of it, and there's

another neuron network that can decompress it. So basically this compressed form can only be read by Chameleon.

Is the neuron network being fed and growing by everyone using it?

No, the training must be done in-house, it cannot be done on the fly. If someone needs something to be changed, we can re-train it and then we have to ship out a new plug-in, but the users don't train the neuron network because we need to control the quality of the audio we give it as otherwise it could destroy its performance.

We have not changed the neuron network since it came out 2 years ago, the first time will be when we introduce the early reflection parameter, which is still in its early stages.

Have you considered giving more control to the users, like decay time, pre delay etc. to make it into kind of a hybrid between convolution and algorithmic reverb?

For now, there's the reverb time and EQ. The idea is to extract the reverb from the audio file, but it could be a good idea to estimate the time and also give the user more control to change it in case it is not satisfactory. I will note this one down, it's a good idea. Maybe we will add more parameters to the early reflection version. Chameleon came out two years ago and Chameleon² came out last year, it had workflow related changes, for automation etc. Chameleon³ will probably have some algorithmic changes for estimation.

Chameleon might have changed how people think of matching ADR since it is quite easy to match reverbs now. Have you received feedback on how it has changed their way of working?

No, this is one of the weak points of Accentize. We are not very close in touch with product users. We very sparsely have feedback, so we need to talk to more people and get in touch regarding the production process. We come from a technical and research background, so it is an important step to merge both worlds. We would like more feedback so we can be in touch with the needs of people today.

Thanks for your time.