

LES VOIX QUI SORTENT DU CADRE

Emilien PICHON

département son, promotion 2025

26 mai 2025

Sous la direction de : Valérie DELOOF et Jean-Pierre LAFORCE

LES VOIX QUI SORTENT DU CADRE

Comment et dans quels buts le cinéma s'émancipe-t-il des normes de la voix humaine ?

INTRODUCTION

I - LES NORMES DE LA VOIX HUMAINE

- 1 - La prosodie**
- 2 - Les normes en conversation**
- 3 - Les normes physiques de la voix parlée**

II - QUITTER LES NORMES DANS LE CINÉMA EN PRISE DE VUE RÉELLE

- 1 - Les normes en prise de vue réelle**
- 2 - Repenser la prosodie**
- 3 - L'altération de la voix humaine**
 - a - Les créatures**
 - b - Personnages à l'aspect humain**
- 4 - Le rapport corps/voix**
 - a - Le synchronisme**
 - b - L'emprunt d'une voix**

III - QUITTER LES NORMES DANS LE CINÉMA D'ANIMATION

- 1 - L'abandon des normes est une norme en soit**
- 2 - Le rapport image/voix dans l'animation**
- 3 - Des procédés singuliers d'utilisation de la voix en animation**
- 4 - Les voix dans mon projet de fin d'études**

CONCLUSION

INTRODUCTION

La voix peuple notre quotidien. C'est notre principal outil pour communiquer avec le monde qui nous entoure. Elle nous permet de créer des liens, d'exprimer nos idées, nos émotions, de partager des informations etc. La voix est donc au centre de la vie de chacun et il n'est donc pas étonnant que cette importance soit retranscrite au cinéma. Depuis l'avènement du parlant, la plupart des films sont considérés comme "vococentristes", c'est-à-dire centrés sur la voix. Cette dernière étant souvent interprétée comme l'élément le plus important de la bande son.

En effet, elle se fait souvent le principal vecteur d'informations et de narration dans le cinéma contemporain. De plus, elle permet de donner vie aux personnages, qui, comme nous, peuvent l'utiliser pour exprimer leurs idées, leurs émotions etc. Elle est très souvent à l'origine de l'empathie que développent les spectateurs pour les personnages, mais également de manière plus large, de l'empathie que développe l'humain pour les autres.

Quel rapport entretiennent alors les voix de cinéma avec celles de notre quotidien ? Se doivent-elles d'être au plus proches de notre réalité pour permettre aux spectateurs d'appréhender le film ? Non. Pour preuve, très nombreux sont les films qui dérogent aux normes vocales de notre réalité. Et c'est à ces cas que je vais m'intéresser dans ce mémoire en me posant la question suivante : Comment et dans quels buts le cinéma s'émancipe-t-il des normes de la voix humaine ?

Pour comprendre "comment" le cinéma quitte les normes, je vais d'abord devoir définir ce que sont ces normes. Ce sera donc le thème de ma première partie : qu'est-ce qui fait qu'une voix humaine est ce qu'elle est ? Je me pencherai d'abord sur la prosodie. Terme qui regroupe tous les caractères de la voix qui ne sont pas liés au sens des mots. Je me concentrerai ici sur la prosodie française (étant la langue que j'utilise au quotidien). Je parlerai ensuite des normes de conversation. Puis je finirai en expliquant les normes physiques de la voix, communes à tous les êtres humains.

Une fois ces standards définis, je pourrai donc parler de cas dans lesquels elles sont mises de côté au cinéma. Je m'intéresserai aux dispositifs utilisés, aux buts que peuvent avoir ces abandons de normes ainsi qu'aux effets qu'ils peuvent générer sur le spectateur. Je m'intéresserai tour à tour à trois moyens de quitter les normes vocales : la remise en question de la prosodie courante, l'altération de la voix humaine par des effets ou autres processus et enfin le rapport entre la voix et le corps qui la porte.

Enfin je me pencherai sur le cas de l'animation. J'expliquerai en quoi ce médium a transformé l'abandon des normes de la voix en une norme à elle-même. Je parlerai ensuite du fait que l'image et le son ne sont pas contraints de garder une distance commune avec notre réalité en animation. Je ferai cela en parlant d'exemples de rapports corps/voix qui me semblent pertinents pour étayer ce propos. Enfin j'évoquerai comment, malgré ces fréquents abandons des standards de la voix, l'animation peut tout de même jouer avec ces derniers à des fins narratives.

I - LES NORMES DE LA VOIX HUMAINE

Les enjeux d'un dialogue ne résident pas seulement dans le contenu même de ce qui est dit : «il y a aussi l'identité des locuteurs que ces paroles mettent en présence.»¹

1 - La prosodie

Pour cette partie, je me base principalement sur *La Musicalité de la Voix Parlée* de Grégory Beller - IRCAM (2005) ainsi que sur *Le dialogue* de Claire Vassé (2003)

Au-delà du sens des mots, la voix se caractérise par sa mélodie, son rythme, sa hauteur et son intensité. Tous ces critères liés au langage mais pas directement au sens même des mots se rassemblent sous un terme : la prosodie.

Celle-ci se manifeste par le jeu simultané de plusieurs paramètres acoustiques : la fréquence fondamentale "Fo", le timbre, l'intensité et la durée des phonèmes. Perceptivement, la hauteur et son évolution, le rythme et le tempo (débit), le registre et

¹ Claire Vassé, *Le dialogue*

le timbre mais aussi les pauses et les silences nous permettent la compréhension d'informations au-delà des mots prononcés. C'est cette deuxième partie du double codage de la parole qui lui confère un caractère "naturel" et évite la monotonie. Elle permet entre autres de véhiculer des informations ectolinguistiques (expressivité) ou phonostylistiques (sentiments).

La prosodie est différente dans chaque langue (elle est d'ailleurs plus facile à interpréter dans une langue que l'on comprend mais peut aussi être d'une grande utilité pour cerner les intentions d'un langage que l'on ne parle pas).

Elle varie selon les personnes et les situations : la diction peut être plus ou moins rapide, haute, monotone, timbrée, forte etc...

Deux termes sont utiles pour exprimer les singularités de prosodie : aux traits de langage propres à un individu (idiolecte) s'ajoutent les caractéristiques propres au groupe social auquel il appartient (sociolecte).

Elle joue des rôles multiples dans le langage de tous les jours. Elle reflète la structure hiérarchique de la phrase. Elle distingue une question d'une réponse. Elle "désambiguïse" des séquences telles que « Je ne veux pas mourir idiot » (la prosodie doit préciser lequel des deux est l'idiot, le locuteur ou l'interlocuteur). Elle exprime des attitudes, des émotions.

Même si elle est très variable, on lui trouve tout de même des normes qui sont suivies (la plupart du temps inconsciemment) par le locuteur. « Bien qu'il existe de nombreuses manières de parler, le domaine de la parole reste bien plus homogène que le domaine musical qui présente une énorme variété de types et de formes différentes. »²

Voici quelques normes qui concernent la langue française :

Le français est une langue « à rythme syllabique », c'est-à-dire que les syllabes ont une durée relativement régulière, sans grande différence de longueur entre syllabes accentuées et inaccentuées.

² Klaus R. Scherer, *Emotion expression in speech music*

L'accent tonique en français se situe sur la dernière syllabe du groupe prosodique. Cet accent final se manifeste notamment par un léger allongement de la dernière syllabe et, selon le contexte, une modification de la hauteur mélodique.

Il existe en français un accent d'insistance (ou accent secondaire) qui peut se placer sur une syllabe non finale pour mettre un mot ou une partie de mot en relief.

Les phrases sont découpées en groupes intonatifs. Chaque groupe a son propre schéma mélodique, souvent marqué par une légère pause avant de passer au groupe suivant. Cet interval temporel aide au traitement de l'information stockée dans la mémoire à court terme de l'auditeur. L'intonation des phrases permet d'indiquer à l'interlocuteur la nature de la phrase et parfois l'émotion dont elle est chargée. Bien que de nombreux modèles co-existent, voici *Les Dix Intonations de base* selon Delattre (1966).

<i>Si ces œufs</i>	Continuation mineure		2-3	C-261 F#-92	<i>si se s s</i>	2-3	C-261 F#-92	<i>- Si ces œufs</i>
<i>étaient frais,</i>	Continuation majeure		2-4	C-261 F#-92	<i>e t e f r s</i>	2-4	C-261 F#-92	<i>étaient frais,</i>
<i>j'en prendrais.</i>	Finalité		2-1	C-261 F#-92	<i>j & p r & d r s</i>	2-1	C-261 F#-92	<i>j'en prendrais.</i>
<i>Qui les vend ?</i>	Interrogation		4-1	C-261 F#-92	<i>k i l e v &</i>	4-1	C-261 F#-92	<i>Qui les vend?</i>
<i>C'est bien toi ?</i>	Question		2-4+	C-261 F#-92	<i>s e b j e t v a</i>	2-4+	C-261 F#-92	<i>C'est bien toi?</i>
<i>Ma jolie ?</i>	Echo		4-4	C-261 F#-92	<i>m a j o l i</i>	4-4	C-261 F#-92	<i>ma jolie?</i>
<i>Evidemment,</i>	Implication		2-4-	C-261 F#-92	<i>e v i d a m &</i>	2-4-	C-261 F#-92	<i>- Evidemment,</i>
<i>Monsieur.</i>	Parenthèse		1-1	C-261 F#-92	<i>m e s j s</i>	1-1	C-261 F#-92	<i>Monsieur.</i>
<i>Allons donc !</i>	Exclamation		4-1	C-261 F#-92	<i>a l s d s</i>	4-1	C-261 F#-92	<i>- Allons donc!</i>
<i>Prouve-le-moi.</i>	Commandement		4-1	C-261 F#-92	<i>p r u v l e m v a</i>	4-1	C-261 F#-92	<i>Prouve-le-moi.</i>

Il est assez intéressant de voir en lisant les phrases d'exemple que l'intonation prise naturellement semble bien correspondre aux courbes définies à côté.

Cette prosodie apparaît naturellement pour quelqu'un dont la première langue est le français. Mais elle peut être moins évidente à assimiler pour quelqu'un qui apprend la langue. Ma mère, qui est professeure de FLE, m'explique qu'elle enseigne la prosodie à certains de ses élèves, notamment chinois et japonais. L'exemple qui lui semble le plus courant est de leur conseiller de monter vers les aigus à la fin de leur phrase quand ils posent une question.

En français, la voix parlée peut se diviser habituellement en 2 facteurs temporels : la voix sonore elle-même 85% et le silence entre les mots 15 %.

Pour résumer, bien que chaque voix soit différente et que chaque personne ait sa propre manière de parler, sa propre prosodie, on trouve tout de même dans la langue française (et dans toutes les autres) des standards qui se sont créés naturellement pour faire du langage un meilleur moyen de communication et d'expression des sentiments.

La prosodie évolue bien sûr aussi avec le temps, l'évolution de la langue. Il suffit de regarder par exemple des archives de L'INA datant des années 70 pour comprendre qu'en cinquante ans, au-delà de la qualité de restitution des enregistrements, la manière la prosodie et le vocabulaire ont drastiquement changé.

Enfin, il me semble impossible de ne pas parler du lien entre la forme amenée par la prosodie et le fond. Notre prosodie est bien évidemment guidée par le contexte dans lequel elle s'inscrit. Prenons par exemple une phrase simple : «je sais pas» (le «ne» étant maintenant rarement prononcé dans le langage courant). Si le locuteur la prononce en étant en colère, frustré par l'insistance de son interlocuteur, il pourra hausser le ton et prononcer chaque mot sèchement, séparé de très brefs silences entre chaque, pour saccader sa phrase. Si le locuteur "sait" en réalité et souhaite le faire comprendre, il pourra ralentir légèrement son débit naturel, lier les mots et monter vers les aigus sur la fin, pour exprimer une certaine malice.

2 - Les normes en conversation

« La voix est un instrument fondamental de la sociabilité humaine : elle est le moyen, le canal le plus usité, le plus courant de toute relation entre plusieurs personnes. Plus facile dans la proximité, elle peut cependant également plus ou moins vaincre les distances en jouant sur son intensité. Dans un dialogue, dans une conversation, elle permet d'entamer, d'entretenir, de mettre fin au rapport que l'on a avec l'autre par l'intermédiaire de la parole. »³

³ Alexandre Charvet, *La voix et ses métamorphoses dans « les métamorphoses » d'Ovide*

Au-delà des normes prosodiques, il existe également des normes de conversation que H. Paul Grice synthétise dans *Logique et conversation* (1975).

«Nos échanges de paroles ne se réduisent pas en temps normal à une suite de remarques décousues, et ne seraient pas rationnels si tel était le cas. Ils sont le résultat, jusqu'à un certain point au moins, d'efforts de coopération. Nous pourrions formuler en première approximation un principe général qu'on attendra à voir respecté par tous les participants : que votre contribution conversationnelle corresponde à ce qui est exigé de vous, au stade atteint par celle-ci, par le but ou la direction acceptés de l'échange parlé dans lequel vous êtes engagés. Ce qu'on pourrait appeler principe de coopération»⁴

Ici, la précision "jusqu'à un certain point" est importante. En effet, un locuteur peut évidemment se défaire de cet effort de coopération, en cas de conflit par exemple. Mais il est en effet facile de considérer que cette idée de coopération est à l'épicentre de ce qui fait une conversation normale. Nos participations à une conversation invitent généralement notre interlocuteur à réagir. Une question attend une réponse, une remarque attend une relance, une plainte attend une excuse ou une justification.

Ces normes sont mises en évidence lorsqu'elles sont brisées dans le sketch de Roman Frayssinet *Conversation* (2020). Ce dernier est en soirée et est bloqué face à un invité très bavard qui lui raconte des anecdotes qui l'ennuient. On rentre alors dans sa tête où deux Roman discutent dans le but de trouver une stratégie pour mettre fin à la conversation (dialogue intérieur).

« Roman 1 : "Viens, on se barre"

Roman 2 : "Non, on peut pas juste partir comme ça"

Roman 1 : "Mais bien sûr qu'on peut, on pivote la hanche, on met le genou, bim, on part"

Roman 2 : "Non mais physiquement on peut, mais la politesse veut qu'on peut pas" »⁵

⁴ Herbert Paul Grice, *Logique et conversation*

⁵ Roman Frayssinet, *Conversation*

Roman 1, pris par la colère, n'a plus aucune envie de respecter les normes de conversation. Roman 2, quant à lui, les rappelle et souhaite les respecter. Il semble évident qu'il est difficile dans une discussion du quotidien de partir sans dire un mot et sans justifier son départ alors que l'on n'est pas en conflit. Le principe de coopération serait alors brisé.

Roman 2 propose alors une stratégie qui respecterait ce principe : «on le regarde droit dans les yeux et on lui dit “excusez-moi monsieur, je dois aller parler à quelqu'un là-bas” »⁶. Mais Roman 1 prend le contrôle et décide d'appliquer une nouvelle idée : alors que son interlocuteur monopolise toujours la conversation, Roman pousse un cri fort et aigu en regardant l'homme dans les yeux. Le principe de coopération est alors brisé, rendant la situation absurde. Le bavard, dérouté, part pour aller parler à quelqu'un d'autre.

Il dira d'ailleurs plus tard dans son spectacle *Ô dedans*⁷ (2025) : «C'est fou quand même que l'endroit le plus dur à quitter sur cette planète, c'est une conversation».

3 - Les normes de la voix parlée

Pour cette partie, je me base sur le livre *La voix parlée* de Guy Cornut.

Lorsque l'on parle, la hauteur tonale oscille autour d'une fréquence moyenne qui est pratiquement toujours identique pour un sujet donné. Fréquence que l'on appelle «Fondamentale usuelle de la parole». Cette tonalité moyenne est différente suivant qu'il s'agit d'un enfant (entre 247Hz et 330Hz), d'une femme (entre 196Hz et 294Hz) et d'un homme (entre 98Hz et 147Hz). Après un petit test au spectrogramme, ma fondamentale usuelle semble être aux alentours de 108 Hz.

À l'intérieur de chacune de ces trois catégories, les différences individuelles sont importantes et dépendent de facteurs anatomo-physiologiques (taille des cordes vocales par exemple), familiaux, sociaux, etc.

⁶ *ibidem*

⁷ Roman Frayssinet, *Ô dedans*

L'intensité de la voix peut varier dans des proportions considérables depuis la voix murmurée presque inaudible jusqu'au cri dont l'intensité peut atteindre des valeurs difficilement tolérables par l'oreille (de l'ordre de 110 décibels).

Le timbre de la voix tel qu'on l'analyse au sortir de la bouche est la résultante de la transformation et du modelage du son laryngé par les cavités de résonance.

On peut admettre que le conduit vocal a un rôle de filtre variable ; chacune de ses configurations donne un spectre acoustique particulier grâce auquel on peut identifier les timbres des voyelles ou timbres vocaliques.

Le timbre vocalique, qui varie suivant chaque voyelle, correspond aux zones formantiques. Il représente donc les traits acoustiques fondamentaux qui sont communs à tous les sujets et qui permettent l'identification des diverses voyelles quelle que soit la personne qui parle.

À l'opposé, on donne le nom de timbre extra-vocalique à la partie du spectre acoustique qui varie non pas en fonction des voyelles mais en fonction de la personne. Ce timbre extra-vocalique, qui est propre à chaque individu, dépend moins des caractéristiques anatomo-fonctionnelles du conduit vocal que du son laryngé initial. La partie aiguë du spectre au-delà de 2 800 Hz ne joue pratiquement pas dans la discrimination des voyelles mais a par contre un rôle prédominant pour la couleur de la voix.

II - QUITTER LES NORMES DANS LE CINÉMA EN PRISE DE VUE RÉELLE

1 - Les normes en prise de vue réelle

Le cinéma tente (dans la plupart des cas) de respecter deux principes en ce qui concerne le dialogue : l'intelligibilité et la crédibilité. L'objectif étant de permettre aux spectateurs de comprendre ce qui est dit mais également de le rendre cohérent dans l'espace du film. Cette question touche tous les métiers du son au cinéma.

L'ingénieur du son porte bien sûr une grande importance à l'intelligibilité de la parole. Selon Yves-Marie Omnès, qui est intervenu à l'école, capter les voix de manière claire et cohérente est le centre même de son activité. Il nous avait montré comme exemple un plan séquence de *Queer* de Luca Guadagnino (2025).

Ce dernier semblait complètement impossible à percher de par la complexité de la chorégraphie et le son des machines. Yves-Marie a donc demandé un son seul des répliques de la séquence in situ et l'a monté lui-même sur protools pour s'assurer de la cohérence du texte à l'image. Je pense par ailleurs avoir reçu des sons seuls de répliques sur l'ensemble des projets sur lesquels j'ai travaillé cette année en tant que monteur son et cela malgré des directs souvent très corrects. C'est, selon moi, une très bonne méthode de travail car, au-delà de sécuriser une réplique sur laquelle on a un doute, elle permet d'obtenir des variations de jeu qui peuvent s'avérer plus intéressantes que le direct.

Le deuxième point est le rapport à l'image. Sur cela aussi, l'ingénieur du son a une grande influence. Le placement de la perche et l'équilibre perche/HF qu'il livre, donne sa direction du rapport au cadre. Je précise "sa direction" car elle est un point de vue et non un simple suivi de normes de distances à respecter. L'ingénieur du son pouvant chercher, avec divers placements de perche plus ou moins distants, ce qui sonne pour lui comme étant au plan. La contrainte du champ dans lequel la perche ne doit pas entrer influence bien sûr ces placements mais, cela ne signifie en aucun cas que la position idéale pour l'ingénieur soit un bord cadre au-dessus de l'acteur. Les sons seuls de répliques évoqués au-dessus permettent d'ailleurs de s'affranchir de cette contrainte.

Cette première proposition est ensuite transférée à l'étape du montage. Le monteur des directs peut retravailler l'équilibre perche/HF. Il possède de nombreux outils pour à nouveau questionner l'intelligibilité et les rapports au plan de la parole (nettoyage, choix de doubles, égalisation, compression etc...).

Le monteur son a bien sûr lui aussi un rôle à jouer sur ces questions. Les effets peuvent avoir une influence importante sur l'intelligibilité de la voix. Gwennolé Le Borgne, qui était mon maître de stage en 3ème année, m'a partagé une méthode liée à cette question

que j'ai appliquée lors de mon stage. Avant de travailler sur les effets d'une séquence dialoguée, il écoute cette dernière en se demandant à quel moment il a de l'espace pour placer des événements sonores (qu'il identifie avec des marqueurs). Cela lui permet de ne pas surcharger la bande son, de choisir des événements pertinents et de respecter l'importance de la voix (là encore, bien sûr, quand c'est nécessaire). L'idée n'est pas de ne rien placer sous la voix mais plutôt de rester conscient de sa présence et de sa place dans la bande son.

Le travail des ambiances a aussi bien sûr une grande influence sur l'intégration de la voix dans l'image. Les paroles et les ambiances partagent l'espace créé par la bande son et s'entraident souvent pour le rendre crédible et cohérent.

Le mixeur, enfin, est le dernier à remettre en question ces deux critères, cette fois-ci dans une optique d'écoute en salle. Il décide, avec le réalisateur, du rapport final entre la voix et le reste de la bande son, entre la voix et l'image. À cette étape, même si elles peuvent déjà avoir été travaillées en amont, les réverbérations sont un outil central. Elles ont de très nombreux intérêts : homogénéiser les voix, travailler avec précision leur intégration dans l'espace du film, intégrer des post-synchronisations etc...

Un outil intéressant qui est fréquemment revenu dans les mix auxquels j'ai participé cette année en tant que monteur son est le plugin "Boom It". Il peut notamment être utilisé pour créer de manière artificielle des détimbrages de perche que l'on pourrait souhaiter entendre quand un comédien se met dos à la caméra, par exemple. Cet effet pouvant d'ailleurs être obtenu dès l'étape du tournage si l'ingénieur du son et son perchman décident de ne pas suivre l'acteur à la perche et d'assumer que l'on perde légèrement sa voix.

Ces deux critères sont donc toujours centraux dans la manière de traiter la voix au cinéma, mais ils peuvent tout de même être approchés de manières infiniment variées. Généralement, grâce à ce que permettent la prise de son, le montage son et le mixage, un film peut choisir son ou ses points d'écoute pour une scène donnée. Ce concept est facilement accepté par le spectateur qui ne sera pas dérouté d'entendre des voix très premier plan, très bien définies portées par des personnages qui sont filmés en plan large. Cela est sans doute lié au fait que ce genre d'effet a été permis très rapidement

dans l'histoire du cinéma parlant. Il est donc très courant que l'on ne décide pas simplement de prendre le point d'écoute de la caméra mais plutôt que l'on décide de prendre un point d'écoute sensible qui sert la narration de la scène. C'est un travail auquel Olaa Zhyzkho est très sensible. Lors de son montage son de TFE *Just Drifting* puis de son mix (avec Louis Bellec) elle a très souvent souhaité modifier la présence des voix à l'image, non pas par souci de réalisme et de distance, mais pour marquer l'importance d'un personnage par rapport à un autre et pour guider l'œil des spectateurs.

2 - Repenser la prosodie

Un passage du livre *Le dialogue* de Claire Vassé (2003) aborde ce sujet. Elle commence par rappeler que «le dialogue au cinéma est classiquement considéré comme un élément qui doit se fondre dans la réalité représentée, laisser l'impression que ça parle ni plus ni moins comme dans la vie ni mieux, ni moins bien.»⁸

Mais elle écrit ensuite sur le “parler faux”. Une manière de parler qui ne correspond pas à ce que l'on rencontre dans la vie réelle. Autrement dit, une manière de parler qui semble quitter les normes de la prosodie que nous avons évoquées juste au-dessus.

Voici ce qu'elle dit de cette expression :

«Parler faux ? Une formule qui soulève des débats épineux, qui sous-entend qu'il y aurait un “parler juste”. Mais selon quels critères ? Et pourquoi parler juste serait une qualité et parler faux, une déficience ? [...] parler juste n'est pas forcément le gage d'un bon dialogue. Tout dépend du film, de son rapport à la réalité, de son degré de stylisation, de sa distance aux personnages. L'objectif n'est pas de parler juste mais de raisonner vrai. D'où l'importance du parler faux chez les cinéastes pour lesquels la représentation du monde ne va pas de soi. Parler faux, c'est refuser de réduire le dialogue à un outil narratif, à un élément participant à l'élaboration d'un climat réaliste, à la construction d'une histoire vraisemblable, à l'expression du personnage comme unité indivisible.»⁹

⁸ Claire Vassé, *Le dialogue*

⁹ Claire Vassé, *Le dialogue*

On peut comprendre via cette citation que «parler faux» et «mal jouer» ne sont pas forcément liés. Qu'aller chercher au-delà des normes de la parole peut être une intention de mise en scène à part entière qui trouve son intérêt ailleurs que dans la retranscription de notre réalité. Dans son livre, elle cite de nombreux exemples : le parler théâtral chez Renoir, le parler très neutre chez Bresson, le parler incantatoire chez Duras.

Je vais d'abord me concentrer sur l'exemple de Bresson.

Le premier film que j'ai regardé de Bresson était *Au Hasard Balthazar* (1966) et je me souviens avoir eu beaucoup de mal à accepter la manière de jouer des acteurs. Il me semblait qu'ils récitaient leur texte et qu'ils avaient interdiction de charger leur voix d'émotions. Comme dit précédemment, la prosodie dans notre quotidien est étroitement liée au contexte et aux émotions mais dans ce film, quel que soit l'enjeu de la scène (une dispute, une déclaration d'amour etc.) le ton reste toujours le même. À certains moments, les personnages poussent un peu plus la voix mais semblent être comme restreints dans le volume qu'ils sont autorisés à atteindre. Cela même quand l'objectif porté par la voix est purement de vaincre la distance pour se faire entendre : par exemple quand le père de Mary cherche sa fille au milieu des champs et qu'il l'appelle avec une voix toujours trop faible «Mary Mary»¹⁰.

Lorsque les personnages du film ont une émotion "trop forte" à exprimer, ils semblent privés de voix. Par exemple lorsque Mary est effrayée lorsqu'on lui touche la main alors qu'elle se croyait seule sur un banc, elle fuit sans qu'un son ne sorte de sa bouche. Il en va de même peu après quand elle voit son âne se faire frapper.

J'avais malgré tout fini par accepter cette diction en pensant que son objectif était de ramener toute l'empathie du spectateur sur l'âne et ses mésaventures (une analyse sûrement très infondée mais qui m'avait semblé intéressante en regardant le film).

Dans son intervention, *Voix et sacrifice dans le cinématographe selon Robert Bresson* au colloque BnF/ARM le 5 décembre 2015, Jérôme Thélote explique que le style de voix Bressonien est né dans *Un condamné à mort s'est échappé* (1956) et qu'il a, dans ce film, une justification scénaristique. La manière de parler de Fontaine, neutre, calme, automatique, est justifiée par le besoin du personnage d'être discret, de se méfier de ses geôliers. D'après lui, le sens même de cette diction, surpasse le sens des mots. C'est la

¹⁰ Robert Bresson, *Au Hasard Balthazar*

diction elle-même qui se donne dans la matière sonore du film. Suite à ce dernier, Bresson imposera cette manière de parler à tous ses acteurs (qu'il appelle modèles) car c'est pour lui celle qui se rapproche le plus du «juste».

Voici un extrait d'entretien dans lequel Bresson parle de cette prosodie, paru dans *Le cinéma selon Bresson* de Pierre Ajame, Les Nouvelles littéraires, 26 mai 1966.

«P.A. — Et la fameuse «diction blanche» que vous leur imposez?

R.B. — D'abord, cette diction n'est pas blanche, elle est vraie, je veux dire: juste. Ce mot, «juste», n'est pas toujours compris au théâtre. La diction, au théâtre, s'appuie sur des sentiments étudiés, des sentiments flottants; elle est flottante. Il faut que les paroles, comme les gestes, soient automatiques. C'est de l'automatisme, qui compose les trois quarts de notre existence, que sort le vrai, et non du pensé et du raisonné. Les paroles sont avant tout un rythme. L'importance des rythmes, leur toute-puissance dominent mon travail. Le seul exercice que je fais faire à mes protagonistes est un exercice de lecture, équivalent aux gammes du pianiste.»¹¹

Dans la suite de sa présentation, Jérôme Thélote nous explique que Bresson considérait la prosodie présente dans ces films, comme la diction par excellence d'une subjectivité allergique au monde. Hétérogène aux prescriptions de celui-ci. Une subjectivité absolue qui requiert une philosophie susceptible d'en accepter la possibilité.

Cette diction n'est donc pas pour lui une manière de créer une étrangeté que l'on peut ressentir facilement à l'écoute de ces films. Au contraire, c'est une manière d'être au plus proche du vrai. Le mot de modèle qu'il utilise pour qualifier ses comédiens n'est pas choisi au hasard. Comme il le dit lui-même : «Être modèle, paraître acteur»¹².

Mais j'aimerais m'arrêter sur les mots de Jérôme Thélote : «requiert une philosophie susceptible d'en accepter la possibilité»¹³. Cette phrase semble donner un prérequis pour appréhender la diction dans les films de Bresson. Ce réalisateur semble avoir créé une manière de parler à part entière qu'il faut pleinement accepter pour entrer dans l'univers du film. Cela principalement parce qu'elle semble éloignée de ce que l'on rencontre dans notre quotidien. Pour reprendre les mots de Claire Vasset, Bresson est donc un bon

¹¹ Pierre Ajame, *Le cinéma selon Bresson*

¹² Robert Bresson, *Notes sur le cinématographe*

¹³ Jérôme Thélote, *Voix et sacrifice dans le cinématographe selon Robert Bresson*

exemple de cinéaste qui refuse de réduire le dialogue à un outil narratif, à un élément participant à l'élaboration d'un climat réaliste. Il crée sa propre réalité au sein de son cinéma et c'est ensuite au spectateur d'accepter ou non d'y rentrer.

Un élément de diction particulièrement étrange que j'ai relevé dans *Au hasard, Balthazar* est la manière qu'ont les personnages de répéter deux fois la même chose, exactement sur le même ton (le «Mary Mary» cité au-dessus en est un très bon exemple). Cela apporte quelque chose de presque robotique et déshumanisant aux modèles, ce qui génère chez moi un léger sentiment de malaise.

Un autre cinéaste qui fait appel à ce motif est David Lynch dans son court métrage *The Grandmother* (1969). Dans le film, les parents de l'enfant sont très déshumanisés par leurs gestes et leur voix. On les voit au début du film en train d'agir comme des animaux, d'émettre des cris très étranges semblables aux aboiements d'un chien. Les deux ne répèteront ensuite qu'un seul mot pour tout le reste du film, un mot que je peine à comprendre mais qui semble être le nom du fils «My»¹⁴ (peut-être Marc ou Mike, mais je n'ai pas réussi à trouver cette information). Ce nom est d'abord répété deux fois par le père qui le crie quand il découvre que son fils a mouillé son lit. Deux fois très proches, exactement sur le même ton menaçant. Puis il est repris plus calmement par la mère du garçon qui l'appelle pour qu'il vienne. Également deux fois, très proches, avec la même diction.

Je trouve que ce procédé est extrêmement efficace pour faire des parents, une menace dans le film et pour les mettre sur le même plan, comme s'ils ne faisaient qu'un. Je ressens ici le même sentiment de malaise entraîné par la déshumanisation des personnages. Ce sentiment est amplifié au milieu du film lorsque les parents observent que leur fils ne mange pas. Le père commence par secouer son fils dans un silence qui semble complètement en décalage avec la scène. Ce vide, que l'on ressent d'ailleurs à plusieurs moments du film, a notamment pour motif l'absence totale de son direct. Toute la bande son ayant été créée en post-production. Puis les deux parents se mettent à crier «My» à répétition, de plus en plus fort. Ce procédé me donne également la sensation d'être dans le point d'écoute de l'enfant, comme si l'agressivité et la méchanceté de ses parents avaient transformé toutes leurs paroles en un cri redondant «My» que l'enfant ne prendrait plus la peine d'écouter. On se retrouve alors ici sur un bon exemple d'isolation totale de la prosodie. Ce que disent les parents n'a pas de sens,

¹⁴ David Lynch, *The Grandmother*

seule leur prosodie est porteuse de ce dernier. C'est par cette dernière que l'on comprend que leurs intentions envers leur fils sont mauvaises.

Un autre exemple de déformation de la prosodie est présent chez Lynch dans *EraserHead* (1977) lors de la scène du repas chez les Xs .

On observe cela dès l'entrée d'Henry dans la pièce à vivre. Le «Hello there» de la mère qui le voit entrer depuis son fauteuil, arrive étrangement tard, sans raison apparente. Puis chaque réplique de leur dialogue est séparée par un silence étonnamment long. Cela donne l'impression que les questions et les réponses sont complètement décorrélées. Voici les écarts entre les répliques au début de la conversation

Entrée d'Henry (9s30) «Hello there !» (2s40) «Hello.» (4s60) «I'm very pleased to meet you.» (5s) «Sit Down» (29s30) «It's Henry, isn't it ?» (3s40) «Yes.» (20s) «Mary tells me you're a very nice fellow. What do you do ?» (3s) «I'm on vacation.» (3s) «What did you do ?»¹⁵

Le père arrive ensuite de la cuisine pour rencontrer Henry, il parle très fort, crie presque, sans que la distance ne puisse le justifier. Un écart d'intensité étrange se crée entre lui et Henry.

Le sens des mots dans ces deux exemples devient alors totalement secondaire, notamment via la trivialité des mots échangés. Le dialogue semble être plus un outil qui soutient l'ambiance malaisante de la scène et, comme dans *The Grandmother*, éloigne les parents de Mary de l'humanité que l'on pourrait leur prêter.

Ce procédé est d'ailleurs nettement amplifié par le contexte dans lequel s'ancrent ces discussions dans la séquence chez le Xs. Les tranches dans lesquelles partent Mary sur le canapé puis, plus tard, sa mère à table n'interrompent ni l'une ni l'autre la conversation en cours. Elles sont seulement remarquées par Henry sans que cela n'affecte drastiquement son comportement : il semble anxieux, tendu tout au long du film.

On retrouve chez Lynch comme chez Bresson la nécessité de rompre complètement avec notre réalité pour entrer dans celle du film. Michel Chion écrit par rapport à cela

¹⁵ David Lynch, *EraserHead*

dans son livre *David Lynch* (2007) : «J'ai voulu entamer l'évocation du cinéma de David Lynch en illustrant sur quoi repose cet univers déconcertant : une logique très particulière nous obligeant à renoncer à toute interprétation a priori des comportements et des faits, qu'ils soient pris séparément ou successivement.»¹⁶

Jacques Tati est un autre très bon exemple du détournement de la prosodie. Lui aussi l'isole. Dans *Playtime* par exemple, le réalisateur utilise les voix pour caractériser ses personnages mais pas (ou du moins très rarement) le sens de ce qu'ils disent. Les phrases sont souvent très courtes, interrompues, mêlées entre elles ou au reste de la bande son. Cela fait qu'il nous reste presque exclusivement l'empreinte vocale des personnages. Celle de Monsieur Hulot, le personnage principal du film, est surtout composée de petits fragments de voix souvent presque inaudibles qui caractérisent son aspect "déboussolé", et son insignifiance dans la grande jungle urbaine qu'il visite et qui semble aller trop vite pour lui. Ces petits fragments participent à ce qui définit le personnage, de la même manière que le fait son costume ou sa démarche.

Remarque

Mis à part ces exemples de jeu sur la prosodie au cinéma, on peut également se demander si la prosodie de notre monde est la même que la prosodie du monde de la fiction. Rappelons ici les mots de Claire Vassé : «le dialogue au cinéma est classiquement considéré comme un élément qui doit se fondre dans la réalité représentée, laisser l'impression que ça parle ni plus ni moins comme dans la vie ni mieux, ni moins bien.»¹⁷ Malgré tout, on peut facilement se faire la remarque que, via l'écriture du dialogue, son apprentissage par les comédiens et le fait qu'ils jouent, les standards de prosodie de notre réalité ne sont souvent pas exactement les mêmes que celles du monde de la fiction. Il est, par exemple, rare qu'un personnage bafouille, hésite longuement, se perde dans ses mots si cela n'a pas d'intérêt direct sur sa caractérisation ou sur la narration.

Dans la réalité, si on pense à une personne que l'on considère comme charismatique dans notre entourage et qu'on l'écoute parler, cette dernière aura forcément de temps en

¹⁶ Michel Chion, *David Lynch*

¹⁷ Claire Vassé, *Le dialogue*

temps des artefacts dans son discours même s'il est éloquent. Ce qui ne sera sans doute pas le cas pour un personnage de fiction charismatique qui déclame un discours fort de sens. Si l'acteur bafouille, se trompe, on peut refaire une prise ou faire appel à la post-synchro. Je pense donc qu'au cinéma on ne parle en général pas ni mieux, ni moins bien que dans la vie. Je dirais que le "mieux" est nettement plus courant que le "moins bien". Mais cela est très facile à accepter pour le spectateur notamment, car cette manière de conception est une norme en fiction.

Le film qui m'a récemment mené à cette réflexion est *Under the skin* de Jonathan Glazer (2013). Ce film de science-fiction mélange de manière très intéressante la fiction et la réalité. On y trouve le personnage principal joué par Scarlett Johansson qui roule dans les rues de Glasgow pour trouver des hommes à piéger. Cette partie du film s'ancre dans la réalité, l'actrice s'est réellement promenée dans la ville avec sa camionnette et a discuté avec des habitants qui n'avaient pas la moindre idée qu'il était en train de jouer dans un film. Et dès le début de cette séquence, alors que je n'avais aucune idée du dispositif du film, j'ai immédiatement compris que les hommes chassés n'étaient pas des acteurs. L'aspect un peu surprenant des choix de cadre m'a sûrement aussi aidé à comprendre cela (les caméras ayant dû être dissimulées pour le dispositif), mais je pense que c'est surtout la manière d'être et de parler des personnes qui m'a amené à penser ça. Notamment leurs hésitations, les contradictions dans leurs phrases, leurs regards perdus etc.

Dans ce cadre, on peut donc considérer que les exemples de détournements que j'ai évoqués plus tôt ne s'appliquent pas seulement aux normes prosodiques de notre réalité, mais également à celle de la fiction elle-même. C'est d'ailleurs principalement ces normes que Bresson semble vouloir fuir : «Il faut que les paroles, comme les gestes, soient automatiques. C'est de l'automatisme, qui compose les trois quarts de notre existence, que sort le vrai.»¹⁸

¹⁸ Robert Bresson, *Notes sur le cinématographe*

3 - Altérer ou non la voix humaine

a - Les créatures

Très nombreux sont les exemples d'altération de la voix humaine au cinéma. Les genres auxquels on peut penser tout de suite dans ce cadre sont le fantastique et la science-fiction.

Dans ces derniers, l'intérêt d'altérer la voix est souvent de la prêter à un être qui n'est pas humain mais qui possède tout de même la parole. L'altération vient alors souvent de la volonté de faire coller la voix à la nature ou à la morphologie du personnage.

Les pixies dans *Harry Potter et la chambre des secrets* de Chris Columbus (2002) ou encore les Chipmunks de *Alvin et les Chipmunks* de Tim Hill (2007) sont de petites créatures dont les voix, pitchées, sont bien au-dessus des normes de hauteur de la voix humaine.

Sylvebarbe dans *Le Seigneur des anneaux : Les Deux Tours* de Peter Jackson (2002), Hulk des *Avengers* de Joss Whedon (2012) ou encore Smaug dans *Le Hobbit : La désolation de Smaug* de Peter Jackson (2013) ont une taille massive qui impose des basses que la voix humaine ne peut pas produire mais qui semblent nécessaires pour rendre le personnage crédible à l'écran.

Le processus de création de la voix de Sylvebarbe, l'arbre géant, était malgré tout principalement physique. Pour obtenir une voix profonde qui résonne dans le corps boisé de la créature, l'équipe a créé un long résonateur en bois de forme sinueuse. Les répliques étaient jouées par John Rhys Davis (l'acteur de Gimli) sur le plateau. Elles étaient plus tard projetées à plusieurs reprises dans le résonateur depuis une enceinte, avec un micro à l'autre extrémité, pour obtenir cette voix profonde et puissante qui semble émaner du grand corps de bois du personnage.

Le processus pour Smaug dans *Le Hobbit*, qui prend pour base la voix de l'acteur Benedict Cumberbatch, est également un cas très intéressant. Voici quelques mots de David Farmer, l'homme qui a travaillé sur la voix du personnage : «Il est utile de commencer par une performance exceptionnelle, que Benedict nous a offerte à

profusion. Sa voix est sans doute la meilleure avec laquelle je n'ai jamais travaillé. L'essentiel résidait dans la juste dose d'ajustement de la hauteur (pitch), associée à d'autres améliorations sur les basses, principalement. Ensuite, il y avait les réverbérations qui donnaient l'impression que le son emplissait l'espace, ainsi que le vocodage de grognements d'alligators pour ajouter une couche supplémentaire de profondeur en dessous. Il y a eu un certain nombre de traitements, dont certains ajoutés de manière assez subtile, mais au final, la somme de tous ces éléments a créé quelque chose que j'espère à la fois naturel et authentique.»¹⁹

Cet aspect authentique a d'ailleurs été recherché depuis la prise de voix. Benedict Cumberbatch jouait avec son corps les mouvements qu'il projetait sur ce personnage de dragon géant. Mouvements qui influençaient sa diction alors qu'il jouait ses répliques.

Cette approche ressemble à celle qui a permis la création de la voix de Fix dans *Le Règne Animal* de Thomas Cailley (2023). Théo Serror, qui a travaillé sur la création de cette voix m'a expliqué leur processus. Il a tout d'abord insisté sur le fait que l'aspect technique de la création de cette voix, bien que poussé, était finalement secondaire dans l'obtention du résultat final. Le principal étant les intentions de mise en scène et leur application au son. En effet, Théo est arrivé très tôt sur le projet. Alors que le scénario et les concepts arts étaient les seuls éléments à observer, on lui a demandé d'expérimenter sur des croisements de voix humaines et animales. Cela permettant à l'équipe de mise en scène de pouvoir se projeter sur diverses possibilités.

Il m'explique que la voix de Fix n'a pas seulement été créée en post production (loin de là) mais qu'elle a vu le jour avant même le tournage. En effet, Tom Mercier, qui interprète Fix, a été formé durant une session d'enregistrement par deux imitateurs d'oiseaux. Cette session a permis à l'acteur d'apprendre des méthodes d'imitation, comme par exemple le cri inspiré (très courant chez les volatiles). Cette session a également été l'occasion d'enregistrer des sons de voix en vue de la post- production. Les méthodes qu'il a appliquées au tournage ont influencé son jeu dans sa manière de parler mais également dans sa gestuelle, ayant ainsi une forte influence sur tout ce qui constitue le personnage. Une séance de post-synchronisation, exclusivement dédiée aux vocalises d'oiseaux, a été organisée après le tournage. Théo m'explique que de

¹⁹ Youtube : SoundWorks Collection - The Sound of The Hobbit

nombreuses méthodes ont été utilisées lors de la post-production, notamment de nombreux samples d'oiseaux et d'imitateurs d'oiseaux ainsi que de l'intelligence artificielle capable de transposer des empreintes sonores de corbeaux, de grues, de sifflets aztèques à la voix de Fix. Malgré tout cela, des événements vocaux provenant de Fix imitant l'oiseau ont été conservés dans le film, parfois tels quels, parfois renforcés. Et même lorsque sa voix est remplacée, cela est facilité par le fait que l'intention du son avait déjà été jouée à l'image par le comédien.

On retrouve ici l'idée de recherche d'authenticité. Il est plus simple de réussir à obtenir le résultat souhaité si la recherche de ce dernier a été engagée dès les prémices du projet.

Je trouve ces trois exemples intéressants car ce sont des personnages qui, malgré leur forme particulière, ont une personnalité et une manière de penser qui les rapprochent d'un humain. Je pense donc que ne pas trop ressentir le traitement sur la voix était un enjeu de taille. Un équilibre très subtil à trouver pour correspondre au personnage tout en gardant sa part d'humanité.

Pour les créatures, on trouve d'ailleurs beaucoup d'exemples qui n'utilisent aucun traitement d'effet sur la voix. Les primates dans *La Planète des Singes* de Franklin Schaffner (1968), le *Bon Gros Géant* de Steven Spielberg (2016), ou encore les créatures du BPRD dans *Hellboy* de Guillermo Del Toro (2004).

Tous ces exemples conservent une voix parfaitement humaine, sans doute pour aider le spectateur à pouvoir s'identifier plus facilement aux personnages. En effet, ces derniers font face, dans leur film respectif, à des problématiques auxquelles on peut tous s'identifier. Conserver leur part d'humanité via l'outil de la voix, en dépit de la logique physique, aide sans doute à cela.

Mais ce procédé peut également avoir l'effet parfaitement inverse. Donner une voix humaine à une créature qui est loin d'être humaine peut générer un grand sentiment de malaise. C'est le cas par exemple dans *Le Festin Nu* de David Cronenberg (1991). Dans ce film, Will, le personnage principal est victime d'hallucinations. L'une d'entre elles est une discussion avec une sorte d'énorme cafard géant à l'aspect terrifiant. Cette

hallucination, qui est la première du film, possède une voix humaine. Cette voix, de par son timbre étrange, granuleux, colle très bien à la créature et facilite la compatibilité corps/voix, mais reste tout de même clairement humaine. L'étrangeté de cette voix est amplifiée par plusieurs éléments. Tout d'abord son délai d'apparition. On ouvre la boîte de l'insecte, il révèle sa tête, puis sort de la boîte pour se poser sur une table face à Will. Les sons qu'il produit sont d'abord ceux que l'on attendrait de son corps : des stridulations massives d'insectes. Puis les deux hommes qui étaient avec Will dans la pièce partent. Et seulement à ce moment-là, l'insecte révèle sa voix, humaine. Son apparition à l'image était déjà une surprise mais le fait de retarder la révélation de sa voix permet d'en créer une autre encore plus marquée juste après. De plus, les deux premiers mots qu'il prononce "William Lee" se font entendre lorsque l'insecte est hors champ. Cela nous invite un instant à tenter de comprendre d'où provient cette voix et à penser qu'un autre homme est dans la salle. Mais une fois que l'insecte est montré à nouveau, il n'y plus de doute possible : la créature a un orifice étrange sur le dos qui lui sert de bouche et dont les "labiales" sont synchronisées à la voix. Cette synchronisation amplifie l'ancrage de la voix au corps de la créature et démultiplie le malaise qu'elle entraîne.

Plus tard dans le film, d'autres créatures se feront également entendre, des aliens anthropomorphes, un mix entre un cafard géant et une machine à écrire. Ces créatures seront toujours doublées par la même personne : Peter Boretski. Cette voix commune aux créatures du film permet de marquer la nature hallucinatoire de leur présence mais donne en plus de cela, la sensation qu'une seule et même entité suit et surveille attentivement les faits et gestes de Will. Cela rend leur présence d'autant plus oppressante.

Mais on peut aussi trouver dans ce film le caractère humanisant que peut offrir une voix à une créature. En effet, lorsque Will ramène une autre cafarde-machine à écrire dans sa chambre, son propre cafard-machine finit par la manger vivante. La mort de cette créature d'abord hors-champ est marquée par ses cris, semblables à ceux d'une femme. On croit donc aisément, tant que la source n'est pas révélée, qu'une humaine est victime de quelque chose de terrible. Mais lorsque la situation est montrée, la scène n'en devient pas pour autant plus légère. On développe une étrange empathie pour la créature qui est

pourtant terrifiante visuellement et cela, à mon sens, pour une seule et unique raison : sa voix.

Certaines altérations de voix sont devenues des codes de cinéma pour des profils de personnages en particulier. Les robots par exemple dont la voix est très souvent accompagnée d'un chorus, effet qui donne instantanément un aspect factice à une voix humaine.

On peut citer en exemple C3PO dans *Stars Wars* de George Lucas (1977), TARS dans *Interstellar* (2014) de Christopher Nolan ou encore JARVIS dans *Iron Man* (2008) de Jon Favreau.

Cet effet permet d'identifier la nature robotique des personnages tout en conservant le timbre naturel de l'acteur qui l'interprète. Cela n'altère donc, à mon avis, que très peu la relation que le spectateur peut entretenir avec le personnage. C3PO en est un très bon exemple. Le jeu dynamique et très expressif d'Anthony Daniels, associé au rôle humoristique que tient le robot, permet de donner vie au personnage auquel il est facile de s'attacher.

Bien évidemment tout code généré par le cinéma n'est jamais inébranlable. En restant sur le thème des robots, j'aimerais parler de deux exemples de films qui ne font appel à aucun effet sur la voix de leur personnage artificielle *Her* de Spike Jonze (2013) et *2001, l'odyssée de l'espace* de Stanley Kubrick (1968).

Dans *Her*, Samantha, l'intelligence artificielle jouée par Scarlett Johansson ne prend vie que par la voix et ne pourrait pas sonner plus humaine. Elle respire, rit et est très expressive. La nature de sa voix est au centre de l'enjeu du film. C'est son aspect terriblement humain qui fait que Théodore finit par tomber amoureux d'elle. Le mix de sa voix dans le film, que l'on entend toujours très proche, très défini, décuple sa présence et son emprise inquiétante sur Théodore (et sur le spectateur). Ce choix de mix est d'ailleurs très bien justifié dans le film : on prend le point d'écoute de Théodore qui est équipé d'écouteurs pour parler à Samantha.

Les choix de mise en scène de cette voix sont d'autant plus intéressants aujourd'hui du fait de l'actualité. Ce film d'anticipation a très bien anticipé : les voix actuelles des IA (qui ne sont pas encore au niveau de celle de Samantha) sont de plus en plus marquées par leur similitude à notre prosodie. J'ai moi-même été récemment frappé en parlant à

ChatGTP, lorsque que j'ai réalisé qu'elle reprenait son souffle et hésitait parfois lorsqu'elle me parlait.

Dans *2001, l'odyssée de l'espace*, HAL 9000 inquiète également mais de manière différente. Contrairement à Samantha, la voix d'HAL est très froide, très austère, mais malgré tout, de part son timbre inaltéré, très humaine. Dans *La voix au cinéma*, Michel Chion écrit ceci pour parler de HAL : «Le principal personnage du film 2001 de Kubrick, c'est l'ordinateur HAL, dont la voix résonne dans tout le vaisseau spatial Discovery. Kubrick nous montre un ou deux de ces "yeux" avec lesquels HAL peut voir : ce n'est qu'un voyant rouge, que l'on imagine répandu partout dans l'astronef à d'innombrables exemplaires. Mais il n'a pas jugé utile d'insister visuellement sur l'œil de HAL, chaque fois que celui-ci intervient ou est censé voir quelque chose. Pourquoi ? Parce que Hal est d'abord une voix , et cela suffit [...] il a un oeil dans sa voix [...] il suffit que HAL habite de cette voix tout l'espace du vaisseau pour qu'on lui suppose ubiquité, omnivoyance, omniscience et de grands pouvoirs.»²⁰

C'est en effet toute la force de ce personnage. On a la sensation qu'aucune information ne peut lui échapper et qu'il est complètement impalpable. Il a aussi d'effrayant, comme Samantha, sa proximité avec l'humanité. On pourrait facilement le voir comme le méchant du film mais il ne possède pas de conscience. Malgré tout, et cela est sans doute presque exclusivement lié à sa voix, on ne peut pas s'empêcher de lui prêter de mauvaises intentions.

C'est lors de sa mort que Hal révèle sa vraie nature de robot. Il commence tout d'abord par s'enfoncer plus en profondeur dans la "vallée de l'étrange" : cette théorie du roboticien japonais Masahiro Mori, selon laquelle plus un robot est similaire à un être humain, plus ses imperfections nous paraissent monstrueuses. HAL s'enfonce dans cette vallée par la voix et plus particulièrement la prosodie. Tout d'abord le fond et la forme de ses propos sont complètement en désaccord : il supplie, a peur de mourir, est en perte totale de contrôle, mais son ton reste neutre, égal, froid. Ensuite, il fait quelque chose que j'ai déjà associé au monde robotique plus tôt : il répète plusieurs fois la même chose exactement sur le même ton. «I'm afraid, I'm afraid [...] I can feel it, I can feel it»²¹.

²⁰ Michel Chion, *La voix au cinéma*

²¹ Stanley Kubrick, *2001, l'odyssée de l'espace*

Puis HAL quitte cette vallée pour se révéler complètement en tant que robot : il se réinitialise. On le comprend une fois de plus via la voix, il récite mécaniquement les phrases qu'il est censé prononcer à son allumage. Puis la nature même de sa voix finit par le trahir. Alors qu'il chante, elle s'effondre lentement vers les basses jusqu'à quitter complètement ce qu'une voix humaine serait capable de produire. Cela donne à sa mort un côté grotesque, presque comique qui génère un contre-point étrange avec la gravité de la situation en cours.

Il semble donc que la question d'altérer ou non la voix humaine est toujours une question centrale dans la conception d'un personnage non humanoïde. Ces altérations peuvent permettre de mieux correspondre aux traits du personnage mais il est important de noter que si l'on souhaite donner un côté humain à la créature, il est intéressant de prendre le plus possible d'éléments qui proviennent de l'interprétation du comédien et d'éviter de trop se reposer sur la post-production. L'extrême de cette idée étant simplement de s'affranchir de toute altération. Mais donner une caractéristique humaine à un personnage non-humain n'a pas forcément pour vocation de créer de l'empathie chez le spectateur. Cela peut également avoir l'effet parfaitement inverse et générer un fort rejet.

b - Personnages à l'aspect humain

Parlons maintenant de l'abandon des normes de la voix chez des personnages à l'aspect humain. Je précise "aspect" car, pour plusieurs des exemples que je vais citer, les personnages ne sont pas considérés dans l'univers du film comme des humains, malgré leurs traits.

L'altération des normes chez un personnage a souvent pour effet de le mettre sur un plan différent (souvent supérieur) des humains qui l'entourent. Un super-héros de l'univers X-MEN tire d'ailleurs ses pouvoirs de sa voix : le Hurlleur, interprété par Caleb Landry Jones dans *X-MEN le commencement* de Matthew Vaughn (2011). Tous ses pouvoirs résident dans le dépassement des normes de la voix humaine, il peut crier beaucoup plus fort, plus aigu qu'un humain. Ses cris dépassent même les normes de la physique, ils

sont supersoniques (dépassent les 344 m/s). Le traitement de ces cris, qui lui permettent notamment de voler dans le film, est intéressant. Au lieu de détruire les oreilles des spectateurs, l'équipe a opté pour des cris aigus, désincarnés, qui sont impossibles à associer au timbre du personnage lorsqu'il parle normalement. On sent l'énergie dans ses cris, mais pas tant le côté strident et fort qu'il pourrait avoir. Ils semblent à l'inverse presque fantomatiques, ce qui nous donne l'impression (et c'est d'ailleurs le cas dans l'univers du film) que les cris du Hurlleur sont à peine accessibles à l'écoute humaine classique. Cela renforce, par l'absence d'information, la singularité du super-héros.

Un autre bon exemple est La Voix, le talent des Soeurs du Bene Gesserit que Paul Atreides, interprété par Timothée Chalamet, apprend à maîtriser dans *Dune* de Denis Villeneuve (2021). Cette voix, qui nécessite une formation rigoureuse, permet à son utilisateur de prendre le contrôle de l'esprit d'autrui. La rareté de son utilisation dans le film et l'impact impressionnant qu'elle a dans la bande son, confère à cette voix un grand aura de puissance.

Le monteur son Theo Green explique que cette force provient de la superposition d'une multitude de voix additionnées à celle de l'acteur. On retrouve ici, l'importance donnée à la conservation de la voix du comédien, qui permet de rattacher facilement tout effet additionnel à son corps.

Les voix nombreuses représentent celles des Soeurs du Bene Gesserit à travers le temps et offrent la sensation de révéler un puits ancestral de connaissance. L'autre élément constitutif de la voix est une basse profonde. Theo Green explique qu'il a trouvé l'effet qu'il souhaitait obtenir dans le travail d'un artiste de reggae dub appelé Lee Scratch Perry. La technique consiste à émettre une très basse fréquence via un subwoofer dans une pièce contenant divers meubles et objets et de prendre le son depuis l'autre côté de la pièce. Cela permet, en plus de la basse, de capter les mouvements, les tremblements qu'elle génère dans l'espace et donc de la rendre beaucoup plus physique et impactante. Lors de sa première tentative visible à l'écran, Paul échoue. Cet échec est exprimé au son par un délai entre les mouvements de ses lèvres et la phrase qu'il prononce avec La Voix. Cela révèle alors les deux composants de cette dernière : le layer de basse est synchrone aux labiales mais n'offre aucune intelligibilité. Cette dernière est révélée juste après quand les multiples voix résonnent dans la pièce. Malgré l'échec, on ressent

déjà le grand potentiel de ce pouvoir. Tous les éléments sont là, il suffit maintenant qu'ils résonnent ensemble.

J'ai moi-même utilisé cette année une altération de voix pour donner à un personnage un caractère surhumain. Dans le TFE de Caroline Favé, *Toxic Underground*. La protagoniste principale, Alwena, quitte sa nature humaine pour devenir un monstre symbolisant la relation toxique qu'elle entretient avec l'autre personnage du film, Allan. Cette transformation se fait d'abord par la voix. Alwena exprime sa colère envers Allan avec la phrase «Ta sœur tu lui touches la cuisse comme ça ?»²². Nous avons altéré sa voix en la superposant trois fois avec des hauteurs différentes et en ajoutant une réverbération inversée. L'objectif étant de rendre soudainement Alwena monstrueuse par la voix avant qu'elle ne le devienne physiquement. Cette phrase initialise immédiatement une hiérarchie entre les deux personnages : Allan n'est plus face à une simple humaine mais à quelque chose qui le dépasse.

L'abandon de normes soniques de la voix chez les personnages humanoïdes a donc la capacité de leur faire perdre instantanément leur humanité et de les placer sur un autre plan que le reste des personnages.

4 - Le rapport entre voix et corps

a - Le synchronisme

La norme de synchronisation entre corps et voix est par contre aujourd'hui un critère sévère de crédibilité dans les films. On l'accepte aujourd'hui seulement dans le cadre d'un effet de mise en scène ou pour un film doublé dans une autre langue, car on est conscient du processus qui se cache derrière. Mais même dans ce cas, notamment en France, un grand effort est mis en place pour faire correspondre au maximum la voix que l'on prête à un personnage et le mouvement de ses lèvres. C'est le rôle de l'adaptateur doublage qui traduit les répliques du film tout en adaptant au mieux sa traduction aux labiales des comédiens de la version d'origine.

²² Caroline Favé, *Toxic Underground*

Cette sévérité vis-à-vis du rapport corps/voix n'a pas toujours existé et s'est installée progressivement dans le cinéma. Cela à des rythmes différents selon les pays. En Italie, par exemple, ce critère est resté secondaire plus longtemps qu'en France. Le plus grand représentant de cette absence de correspondance était Fellini. Il décrit son point de vue sur la question dans son livre *Fellini par Fellini* (1984) : « Moi j'éprouve le besoin de donner au son la même importance qu'à l'image, de créer une sorte de polyphonie. Et c'est pourquoi, bien souvent, je suis contre l'utilisation de la voix du visage et de la voix du même comédien. Ce qui m'importe c'est que le personnage ait une voix qui le rende plus expressif. Pour moi le doublage est indispensable, c'est une opération musicale par laquelle je renforce la signification du figuratif. »²³

Aujourd'hui un film qui, dans sa langue originale, ne respecte pas le synchronisme serait sûrement difficilement accepté par le public qui pourrait facilement l'interpréter comme un problème technique.

Alors que la tradition de la synchronisation était déjà très bien installée en France, Marguerite Duras la remet nettement en question dans *India Song*.

Dans ce film de 1975, la réalisatrice crée une désynchronisation totale entre le film des voix et le film des images. En effet, les personnages ne parlent jamais en son synchrone in. Les deux films qui se déroulent en parallèle racontent la même histoire, le souvenir d'une femme disparue appelée Anne-Marie Stretter. Mais les deux le font à un rythme différent, prenant souvent de l'autonomie l'un par rapport à l'autre.

Marguerite Duras décrira son futur procédé dans son adaptation d'*India Song* paru en 1973 : «Ainsi, aucun son n'est présent à l'image, la voix des acteurs ne s'articule pas sur les visages, et les lieux évoqués par la bande-son ne sont pas ceux que l'on voit. L'intrigue ne se situe pas dans l'image, mais dans le son hors-champ. *India Song* se bâtira par le SON, et puis par la lumière»²⁴.

Elle utilisera le terme de «vissage» pour parler du synchronisme mécanique présent au cinéma. Mécanisme avec lequel elle souhaite rompre. Michel Chion nous rapporte sa

²³ Federico Fellini, *Fellini par Fellini*

²⁴ Marguerite Duras, *India Song*

réflexion dans *La voix au cinéma* : « Le cinéma courant, dit Marguerite Duras, veut à tout prix que les voix soient vissées sur les corps, et c'est avec ce vissage, qui est pour elle une tricherie, qu'elle a voulu rompre avec *India Song*, en dévissant les voix et en les laissant errer. [...] Cette idéologie du vissage, avec ses exigences obsessionnelles de synchronisation, on la trouve dans la tradition du cinéma français, attaché plus que d'autres à une synchronisation indécélable.»²⁵

De plus ce procédé s'associe avec une prosodie particulière que la réalisatrice décrit également dans son livre : «La diction, en général, devra être d'une extrême précision. Elle devra ne pas paraître TOUT A FAIT NATURELLE. Un défaut, léger, devra être mis au point lors des répétitions : il sera COMMUN à toutes les voix. On devrait avoir le sentiment d'une lecture, mais rapportée, c'est-à-dire : DEJA JOUEE. C'est ce que nous appelons : voix de lecture intérieure.»²⁶

Dans ce film, l'indépendance totale du son vis à vis de l'image, associée à cette prosodie étrange, semblant souvent désincarnée, permet de ressentir que les événements du film sont passés, qu'ils sont de l'ordre du souvenir : impalpable, décousu, déformé.

Bien plus récemment, dans le thriller *En Décalage* de Juanjo Giménez Peña (2021), le décalage son/image est au centre même de l'intrigue. Dans ce film, une ingénieure du son perd de plus en plus le synchronisme entre ce qu'elle voit et ce qu'elle entend. Elle doit donc quitter son travail et tenter de vivre avec cette étrange malédiction. Le film offre dans ce contexte des scènes dans lesquelles la voix arrive d'abord légèrement après le mouvement des lèvres puis, plus tard, des dialogues complets qui se déroulent sans aucun son de voix, générant une ambiance très étrange. On découvre ensuite le contenu de la discussion après qu'elle est terminée.

Cet effet de désynchronisation corps/voix peut entraîner un sentiment de malaise plus grand encore s'il est ponctuel et complètement inattendu.

C'est ce qui a lieu dans *Le Festin Nu* de Cronenberg. Alors qu'il marche dans la rue, Will est alpagué par le personnage de Tom Frost (interprété par Ian Holm). Après une

²⁵ Michel Chion, *La voix au cinéma*

²⁶ Marguerite Duras, *India Song*

discussion triviale sur les machines à écrire qu'ils utilisent, la conversation biffurque soudainement sur le fait que Will ait tué sa femme (au début du film). Il s'en défend en indiquant que c'était un accident. Mais Tom, lui répond qu'une telle chose n'existe pas, et qu'il est lui même en train de tuer sa femme à petit feu inconsciemment. Will relève le paradoxe : ça ne peut pas être inconscient si on est en train de le dire. Tom lui explique alors que cette discussion n'a pas lieu sur le plan conscient mais inconscient. Pour le prouver à Will (et aux spectateurs), il lui demande de regarder ses lèvres et d'observer qu'elles sont en train de dire autre chose que ses paroles. Une fois ce "fait" établi, il se lance dans l'explication des diverses méthodes qu'il applique inconsciemment pour tuer sa femme. L'étrangeté de cette scène, au-delà du décalage image/voix, est appuyée par la soudaineté de la disparition et réapparition de la synchro. Juste après la phrase «si tu regardes attentivement mes lèvres», on perd complètement la synchronisation labiale et on la récupère aussi soudainement lorsque Tom a terminé de décrire son plan machiavélique. On retombe alors dans une conversation triviale, en ayant l'impression d'en avoir manqué un bout.

Ce procédé développe l'aspect hallucinatoire et imprévisible du film. Il nous invite à rester sur nos gardes. De plus, c'est un très bon moyen d'exprimer le subconscient du personnage. Il prononce ce discours mais ses mots ne sont pas liés à son corps, ils sont complètement inconnus de lui.

La désynchronisation entre l'image et la voix représente donc aujourd'hui toujours un parti pris marqué de mise en scène, un effet qui sera facilement interprété par le spectateur comme un abandon du concret, du réaliste. Cela est étroitement lié à l'idée de norme : maintenant que le synchronisme est un standard au cinéma, il semble difficile de le quitter sans que cela ne déstabilise le spectateur.

b - L'emprunt d'une voix

En 1929, durant le tournage de son film *Blackmail*, Hitchcock est confronté à l'apparition du parlant aux Etats-Unis. Souhaitant, lui aussi, utiliser cette technique, il décide de faire deux versions de son film, une muette, une parlante. Mais un problème s'impose pour la version parlante : l'actrice principale, Anny Ondra, a un fort accent d'Europe centrale l'empêchant, pour Hitchcock, d'assurer les parties parlées. Pour

pallier cela, il fait appel à l'actrice Joan Barry dont la voix sera prêtée au corps d'Anny. Barry récitait le dialogue et était prise au son, tandis que Ondra tentait de "mimer" sa parole devant la caméra. Ce film, qui fut le premier du cinéma parlant anglais, ancrerait donc la voix d'une comédienne dans le corps d'une autre.

Le passage au parlant fut pour certains acteurs un sombre événement. Eux qui avaient développé leur talent autour de la gestuelle, des expressions faciales n'avaient pour autant pas forcément la même aisance pour l'exercice de la parole.

C'est le thème central de *Singin' In The Rain* de Gene Kelly et Stanley Donen (1953). Dans ce film, le prêt de voix d'une comédienne à une autre est le centre même de l'intrigue. Lina Lamont (Jean Hagen), une actrice star du cinéma va être à l'affiche de son premier film parlant. Problème : elle a une voix nasillarde insupportable. Il est donc décidé que sa voix dans le film ne sera pas la sienne (situation très proche de *Blackmail*, malgré des contextes différents). Kathy Selden (Debbie Reynolds) sera donc sa voix. Lina, découvrant cette idée, fait pression sur le studio pour que Kathy soit obligée de la doubler en secret sur tous ses films, lui empêchant ainsi l'accès à la célébrité.

Dans ce cas-là, le spectateur est conscient du subterfuge. Il en est complice et s'inquiète au cours du film du destin réservé à Kelly. Cette dernière finit enfin par récupérer sa voix et à en priver Lina, par la même occasion. À la fin du film, cette dernière interprète *Singin' In The Rain* devant un salle comble, tandis que Kelly lui prête une fois de plus sa voix, dissimulée derrière un rideau. Mais ce rideau se lève et aligne la compréhension du public dans le film à celle du public du film. Toute la salle rit, mais la prestation continue tout de même, transformant Lina, la grande chanteuse, en une figure burlesque, en une impositrice. L'absurdité du moment est amplifiée par l'arrivée de Cosmo Brown (Donald O'Connor) sur scène qui chante quelques mots dans le micro, donnant alors à Lina une voix d'homme durant quelques notes. Cette dernière réalise alors qu'elle est démasquée.

Le fait de prêter une voix qui n'est pas la sienne à un comédien peut également être un effet volontaire pour générer un certain malaise chez le spectateur. *L'exorciste* de William Friedkin (1973) en est, bien sûr, un très bon exemple.

Avant mes recherches sur le sujet, j'étais à peu près sûr que la voix de Regan MacNeil (Linda Blair), la fille victime de la possession du diable de l'exorciste, était un parfait exemple pour parler d'altération de voix. Mais cette voix au ton rauque, granuleux, glaçant est en réalité le résultat d'une grande performance de l'actrice Mercedes McCambridge. Cette voix correspond également à une sortie des normes vocales car la fondamentale usuelle de l'actrice est inférieure à ce que pourrait produire la voix d'un enfant. Mais c'est, malgré tout, une voix humaine authentique.

Le procédé pour arriver à ce résultat est assez drastique et difficile à imaginer pour la production d'un film de nos jours. Ce dernier est décrit dans le documentaire *The Fear Of God - 25 Years Of The Exorcist* de Nick Freand Jones (1998). Mercedes demande à manger des œufs crus, à fumer constamment et à boire pour être complètement relâchée. L'actrice explique que l'équipe de mise en scène l'a attachée sur une chaise au niveau du cou, des bras, des genoux et des pieds avec des morceaux de draps, pour qu'elle se sente comme Reagan dans les scènes où elle est contrainte.

Voilà comment le réalisateur parle de ses souvenirs d'enregistrements «Des choses étranges avaient lieu dans sa gorge. Des sons doubles, triples émergeaient en même temps, des sons de sifflement. Cela était proche de ce que l'on pourrait imaginer d'une personne possédée par divers démons.»²⁷

Comme pour Fix, on retrouve ici l'exemple d'une comédienne qui a interprété au plus proche, avec ce que lui permettait sa voix, une créature non humaine.

Cette association étrange de voix et de corps, liée à l'aspect de plus en plus déformé de la petite fille, lui confère un aspect complètement désincarné. Son corps, qui ne correspond plus du tout à sa voix ou à ses propos, n'est plus le sien. C'est devenu l'habitat, le vaisseau du démon. Le passage à cet état a lieu hors champ : alors que sa mère est au rez-de-chaussée, elle entend au-dessus d'elle, venant de la chambre de sa fille, deux voix. Celle de Reagan et celle du démon. Depuis le hors-champ, on ressent la présence de deux corps dans la pièce. Mais c'est en réalité deux voix qui se battent pour le même corps.

Arrivée en haut, la mère ouvre la porte de la chambre de sa fille. Cette dernière, toujours hors-champ alors que l'on voit des objets s'écraser violemment contre le mur, pousse un dernier cri de sa voix d'enfant. Cri qui sera le dernier avant que la créature ne prenne tout le contrôle.

²⁷ Nick Freand Jones, *The Fear Of God - 25 Years Of The Exorcist*

Donner une voix d'adulte à un enfant ou inversement (en quittant les normes physiques de la voix) permet finalement d'obtenir des effets similaires à ceux de l'altération de voix d'un personnage humanoïde. Le fait que l'on sache qu'un tel corps ne peut pas produire un tel son entraîne immédiatement la perte de la nature humaine du personnage. On le place sur un autre plan et il est facile d'imaginer que sa voix n'est pas le seul élément qui le différencie d'un humain.

III - QUITTER LES NORMES DANS LE CINÉMA D'ANIMATION

Remarque

Le monde du cinéma d'animation est infiniment vaste, aussi bien sur les thèmes abordés que sur les méthodes utilisées, les styles graphiques, la proximité avec notre propre réalité etc. Mon analyse ici va donc se baser principalement sur les films/courts-métrages/vidéos que j'ai vus et sera forcément influencée par ma subjectivité.

1 - L'abandon des normes est une norme en soit

Le cinéma d'animation permet de s'affranchir des standards liés à la voix de manière souvent plus simple sans que cela ne surprenne forcément le spectateur. L'animation a d'ailleurs la particularité de ne pas avoir de lien direct entre le corps et les voix des personnages (sauf quelques exceptions). Les deux sont toujours créés l'un en parallèle de l'autre, partant parfois du dessin pour trouver la voix ou inversement.

La plus grande facilité de dépassement des normes est bien sûr due au rapport à l'image. Cette dernière n'est pas une captation de notre monde mais une reproduction plus ou moins proche de ce dernier ou même un affranchissement total.

C'est cette distance plus ou moins grande avec notre monde qui permet donc de s'éloigner plus facilement de nos habitudes. Moins nous avons de repères dans l'univers que l'on nous propose, moins on peut le juger comme si c'était le nôtre.

L'originalité dans l'utilisation de la voix est donc généralement beaucoup plus permissive dans le cinéma d'animation en particulier lorsque son univers est éloigné du nôtre.

Cela concerne notamment la manière de parler. Des personnages star de l'animation comme Mickey, Homer et Marge Simpson, Bob l'éponge, Wallace, Sammy etc. ont tous une manière de parler et un timbre de voix très caractérisés qui invitent d'ailleurs facilement à l'imitation et que l'on ne pourrait pas du tout imaginer dans notre réalité. Chacun de ces personnages a une identité visuelle très marquée, associée donc à une identité vocale très marquée aussi, l'une sans doute aussi éloignée de notre réalité que l'autre. Leurs réactions, leurs émotions sont souvent plus exagérées que dans le cinéma en prise de vue réelle, et ce notamment parce que l'image peut suivre l'exagération en question (ou inversement).

Les adaptations animées des shonens (catégorie de mangas destinée traditionnellement aux garçons adolescents) comme *One Piece* de Eiichiro Oda ou *JoJo's Bizarre Adventure* de Hirohiko Araki adoptent comme norme prosodique, des manières de parler souvent très intenses, très exagérées, là encore en suivant l'image qui permet d'exacerber drastiquement les émotions. C'est un style d'animation qui semble avoir une prosodie qui lui est propre.

Les transformations vocales, elles aussi, peuvent facilement devenir normales dans l'animation. On peut citer comme exemple *South Park* de Trey Parker et Matt Stone, série dans laquelle les enfants sont joués par des adultes avec une voix pitchée vers les aigus, ce qui est étrangement très simple à accepter pour le spectateur. Trey Parker joue d'ailleurs lui-même quatre des personnages principaux : Stan, Eric, Randy et Mr. Garrison. Matt Stone, lui, en joue trois, Kyle, Kenny et Butters. Les deux jouent de très nombreux personnages secondaires.

April Stewart est, quant à elle, une comédienne responsable à elle seule de la voix de sept personnages féminins récurrents de la série.

Ce genre de cas est très courant en animation. Ainsi Justin Roiland, créateur de *Rick et Morty* (avec Dan Harmon) double lui même ses deux personnages principaux. La

différence d'âge (Rick étant le grand père de Morty) facilite la distinction des deux voix. Deux voix qu'il avait d'ailleurs beaucoup pratiquées sur d'autres projets, longtemps avant la naissance de la série.

2 - Le rapport image/voix dans l'animation

J'ai parlé du fait que s'éloigner de notre réalité était plus simple en animation. Mais il n'y a en aucun cas, une nécessité de s'éloigner autant par l'image que par le son. Les deux ne sont pas forcément liés sur ce critère. C'est très souvent le cas des films Disney et Disney Pixar.

Pour prendre un exemple concret, *Le Roi Lion*, film de Rob Minkoff et Roger Allers (1994) met en scène des animaux de la jungle en animation 2D, mais cela n'empêche pas aux acteurs de jouer avec leur voix naturelle des émotions que l'on aurait sans doute, en fermant les yeux, pu calquer dans notre imaginaire à des gens réels.

Bojack Horseman, série d'animation de Raphael Bob-Waksberg est aussi un bon exemple de cela. Le pitch même de la série pourrait facilement être imaginé en prise de vue réelle : une star d'Hollywood, dont le succès est passée, lutte pour continuer d'exister et résister à ses diverses addictions. La série aborde des sujets très lourds, comme la dépendance, la dépression, la haine de soi. Seulement, les personnages de cette série sont, pour la plupart, des animaux animés dans un style 2D numérique assez statique. Cela n'empêche cependant pas du tout les acteurs de jouer, là aussi, avec une voix naturelle, cherchant l'expression juste, proche de notre réalité.

Les animaux dans cette série ont souvent une personnalité proche de celle que l'on pourrait leur associer : Princess Carolyn, une chatte, est réfléchie et calculatrice. Mr. Peanutbutter, un labrador, est gentil et spontané. Lenny Turteltaub, une tortue, est lent dans ses actions. Ces éléments permettent de bien caractériser les personnages et de faire de certains d'entre eux des figures plus "cartoonesques" que d'autres avec une prosodie plus spécifique et une palette moins large.

C'est aussi, selon moi, ce qui fait la grande force de cette série : elle réussit à passer de moments très drôles, très absurdes et légers à des moments profondément touchants et très lourds sans que l'on ait une réelle sensation de bascule.

Le Roi Lion utilise d'ailleurs aussi la "cartoonisation" de personnages à son avantage en se permettant des scènes humoristiques avec notamment Timon et Pumba.

Dans *Linda veut du poulet !* de Chiara Malta et Sébastien Laudenbach (2023) l'idée d'associer des voix naturalistes à des personnages aux traits simples est très poussée.

Voici quelques mots de la réalisatrice pour parler de son projet : «On voulait vraiment que quelque chose se passe au niveau du jeu, des rencontres entre les comédiens et le scénario. [...] Ce qui nous intéressait c'était vraiment le jeu des comédiens. Voir comment il se déplaçait dans les lieux. Ça nous a vraiment inspirés pour le découpage, pour la mise en scène.»²⁸

Ainsi avec l'aide de son chef opérateur son Erwan Kerzanet, ils ont mis en place des tournages purement sonores, parfois en studio (dont la taille permettait une bonne liberté de mouvement), plus souvent en décor réel. Les comédiens, perchés, pouvaient donc réellement jouer la scène dans un contexte très concret. Cela permettait donc, en plus de capter les voix, d'obtenir le mouvement des corps des comédiens dans l'espace, ainsi que l'influence de leur mouvement sur leur voix. Des accessoires en mousse ont été créés pour le tournage dans le but de permettre aux acteurs de jouer avec, sans pour autant trop parasiter les directs. Un effet de réel indéniable habite donc ce film.

Un autre rapport image/voix très intéressant en animation et qui relève des questions bien différentes est *La Planète Sauvage* de René Laloux (1973).

Dans ce film, on suit les aventures de Terr un garçon «Om» qui, suite à la mort de sa mère, va être adopté par une enfant «Dragg». Les Draggs sont d'immenses créatures humanoïdes bleues qui vivent sur la planète Ygam. Les Oms sont pour eux, de petites créatures à l'intelligence limitée, certains domestiqués, d'autres sauvages. Les sauvages sont pour eux des nuisibles qui envahissent leurs parcs et volent leur biens. Les Oms représentent dans ce film la race humaine qui, sur Ygam, est loin d'être au-dessus des autres espèces.

Je trouve l'utilisation de la voix dans ce film très intéressante. Tout d'abord les Oms et les Draggs ont tous des voix humaines. Cela permet de mieux accepter les Draggs et

²⁸ Youtube : Making-of Linda veut du poulet : la prise de son

leur point de vue, leurs actions. Ainsi on suit le point de vue des deux communautés de manière équivalente (alors que celui des Oms peut paraître plus instinctif à adopter).

Une autre chose qui aide à cela, à mon sens, est le peu de présence vocale des Oms quand les Draggs les attaquent.

Dans la première scène du film, nous commençons à l'échelle humaine. On observe une femme avec un bébé dans les bras qui tente d'échapper à de gigantesques mains bleues (celles des Draggs). On entend tout au long de sa fuite, sur un plan serré, sa respiration haletante. Cette dernière, qui ne révèle pas le timbre de la femme, devient rapidement inaudible. On n'aura ensuite plus la moindre trace de sa voix, ou de celle de son fils dans la séquence. Le mal-être de la femme est exprimé au son par la musique, dont le synchronisme avec les événements de la scène se rapproche du Mickeymousing (technique de musique de film qui souligne chaque événement par la bande sonore).

Une main attrape la femme et la fait tomber de trop haut pour qu'elle survive. Son bébé, toujours complètement muet, s'approche du corps de sa mère.

La violence et la gravité de cette scène, bien que très grandes, est en partie allégée par l'absence de cris et aussi par le biais de l'animation. On pourrait imaginer la même scène en prise de vue réelle, avec les plaintes marquées au son de la mère et de son bébé. Cette dernière serait sans doute extrêmement dure à regarder et diaboliserait immédiatement les Draggs dans l'esprit des spectateurs. Or le film ne peut pas se le permettre pour que l'on continue à accepter les deux points de vue qu'il présente.

Ce processus est d'ailleurs repris lorsque les Draggs procèdent à une "désomisation" du parc dans lequel une communauté d'hommes s'est installée : on assiste à une extermination totale des Oms sans percevoir aucune de leur voix.

3 - Des procédés singuliers d'utilisation de la voix en animation

Comme dit précédemment, l'animation peut se permettre de quitter les normes de la voix humaine sans que cela ne surprenne les spectateurs. Pour autant cela ne veut pas du tout dire que l'animation ne peut pas surprendre avec les normes de la voix.

Un bon exemple de cela est *Anomalisa* de Charlie Kaufman et Duke Johnson (2015). Dans ce film réalisé entièrement en stop motion de marionnettes, on suit Michael Stone, un auteur respecté, mais qui souffre profondément de la monotonie de sa vie. Cette

monotonie est exprimée de manière originale : tout le monde autour de lui à le même visage et la même voix. Mis à part cela, le film est très ancré dans notre réalité aussi bien au son qu'à l'image. Tout l'univers est donc facile à accepter à l'exception bien sûr de cet étrange décalage. En particulier quand cette voix et ce visage (tous deux semblant avoir pour but de représenter un "monsieur tout le monde" d'une quarantaine d'années) sont prêtés à des enfants et à des femmes. Cette étrangeté est aussi multipliée quand beaucoup de gens parlent en même temps, créant ainsi une grande cacophonie d'une seule et même voix. L'élément déclencheur du film est la perception par Michael d'une autre voix. Voix de femme dont il s'empresse de chercher la source. Il finira par trouver Lisa qui est une grande fan de son dernier livre. Ils tombent immédiatement amoureux et entament une relation. Mais cette dernière étant très précoce, Michael finit par se rendre compte que Lisa n'est pas faite pour lui et qu'ils ont peu de choses en commun. Lorsqu'il le réalise, la voix unique de Lisa commence à se mêler légèrement à celle de "monsieur tout le monde". Elle parle avec deux voix, la sienne au premier plan et l'autre sous-mixée. On est ici, comme dans l'ensemble du film dans le point d'écoute de Michael : il est le seul à l'entendre, Lisa ne s'en rend pas compte. Puis soudainement, à sa dernière phrase la bascule se fait : elle a la même voix et le même visage que tout le monde. Michael retombe dans sa monotonie. L'espoir immense que représentait Lisa pour lui s'est effacé.

Bill Wurtz, un artiste et musicien très prolifique sur youtube, expérimente également avec la voix dans l'animation. Il utilise exclusivement sa voix qui a un timbre très doux. Dans *2 friends on the beach* par exemple, il met en scène deux amis qui passent la journée ensemble et discutent. Le style graphique est très particulier : une 3D low-poly avec des textures très peu définies. Les personnages ont des visages extrêmement simplistes et inexpressifs. Un gribouilli monochrome flotte au-dessus de leur tête.

Les deux amis (ainsi que les personnes qu'ils croisent lors de leur journée) ont tous la voix de Bill Wurtz, qui ne tente pas du tout de changer de timbre pour clarifier les sources. Seul un léger pan aide à cela, la voix étant toujours très au premier plan. Il garde un ton très neutre, très plat qui se révèle en particulier lorsqu'un des deux amis rit simplement en prononçant à maintes reprises, toujours sur le même ton, le mot "ha". La vidéo est très bavarde, les dialogues y sont complètement absurdes. Ils sont récités toujours sur le même ton, par la même voix, dans un rythme étrangement rapide par

rapport à une conversation classique. J'ai la sensation, d'une certaine manière, que Bill Wurtz utilise ici le procédé inverse de celui de Jacques Tati dans *Playtime* : il n'isole pas la prosodie, mais plutôt le texte, le fond. La grande monotonie vocale des dialogues nous invite à nous pencher sur ce qui est dit, car c'est, à l'écoute, la seule chose que l'on peut faire. Le ton très sobre et le rythme rapide offrent au dialogue un aspect encore plus particulier. On pourrait presque avoir l'impression de regarder vivre deux aliens qui font semblant d'être humains.

La bande son est constituée de seulement deux éléments, la voix de Bill Wurtz et des virgules musicales qui ponctuent soit les actions des personnages, soit parfois leur parole. Ces virgules entrent, elles aussi, en dialogue avec la voix de Bill Wurtz ce qui offre à l'ensemble de la bande son un caractère musical.

Un autre créateur très présent sur youtube, et dont j'aime beaucoup le travail est David Firth. Dans plusieurs de ses courts-métrages, le vidéaste utilise une technique d'animation digitale en partant de photo-montage provenant d'images réelles. C'est le cas de *Cream* et de *Gice*. Les personnages de ces courts-métrages, aux traits étranges et déformés, ressemblent tout de même, de manière assez déroutante, à des humains. David Firth joue donc ici avec la vallée de l'étrange : la simple vision de ces personnages crée chez moi un léger sentiment de malaise. Sentiment exacerbé par l'usage de la voix.

Cream est un produit révolutionnaire : une crème qui améliore tout ce qu'elle touche, quelle que soit sa nature. L'homme qui présente le produit (joué par David Firth lui-même), le fait sur un ton très calme, très posé. Il parle lentement pour bien être compris. Son ton ressemble à celui d'un gourou de secte. Sentiment intensifié par la présence de nappes de musique relaxantes au début du film. Il devient voix off lorsqu'il explique les bienfaits de la *Cream* prenant alors une place de narrateur, semblant presque omniscient.

Plus tard dans la vidéo, on assiste à une pub pour *Cream*, dans laquelle trois utilisations différentes de la voix mènent toutes au même résultat : un sentiment de malaise.

Premièrement, un homme assis sur un canapé vient de mettre au point un robot humanoïde. Ce dernier a la voix légèrement modifiée par un chorus, un timbre désincarné et le regard dans le vide. Il est inquiétant et semble représenter ce que

pourrait être la vallée de l'étrange au sein même de l'univers du film. Deuxièmement, la fille de l'homme s'approche pour lui dire qu'elle ne se sent pas bien. Sa voix est une voix d'adulte pitchée vers le haut. Dans *South Park*, évoqué plus tôt, ce type de pitch était facile à accepter. Mais dans le cadre de *Cream*, avec le physique déroutant des personnages, cette association corps/voix devient source d'étrangeté, de malaise. Malaise empiré lorsque son père lui administre Cream : elle prononce alors la phrase «non seulement je me sens mieux mais je sens que je suis une meilleure personne»²⁹ avec une voix dédoublée et une réverbération donnant l'impression qu'elle a atteint un stade de connaissance profonde (semblable d'une certaine manière à la sensation de La Voix de Dune). Enfin, le père exprime, d'un ton triste, ce que peut facilement penser le spectateur à ce stade : «Cream a privé la vie de sa magie»³⁰. Il reçoit alors une dose de la fameuse crème puis dit «Hourra, mon enthousiasme est de retour»³¹ sur un ton neutre, faux, aux antipodes de la phrase. Ces deux derniers exemples donnent l'impression que Cream enlève sa nature à l'homme. Cela les rend, elle et ses créateurs, d'autant plus inquiétants.

Nous avons vu que l'abandon des normes de la voix humaine pouvait conférer à un personnage une certaine supériorité. C'est le cas dans *Mind Game* de Masaaki Yuasa (2004). Au début de ce film, Nishi, le personnage principal, meurt, tué par un Yakuza. Son âme se sépare de son corps et arrive dans les limbes où il fait la rencontre de Dieu. Le traitement du personnage de Dieu est très intéressant dans ce film et tire profit des outils de l'animation, aussi bien à l'image qu'au son.

En effet, Dieu est très présent dans la séquence, visible, bavard. Manière assez rare de le représenter. Seulement, il ne cesse de changer de corps et, par la même occasion, de voix et de prosodie. Dieu explique ce phénomène à Nishi «Tu n'arrives pas à décider à quoi je devrais ressembler»³². On comprend alors que l'on partage les points de vue et d'écoute de Nishi. Dieu prend le corps et la voix qu'il lui imagine et ce dernier n'arrive pas à se décider.

On est donc face à un personnage très paradoxal : on le voit et l'entend beaucoup, mais il reste pour autant très mystérieux, insaisissable car on est conscient qu'aucun des

²⁹ David Firth, *Cream*

³⁰ *ibidem*

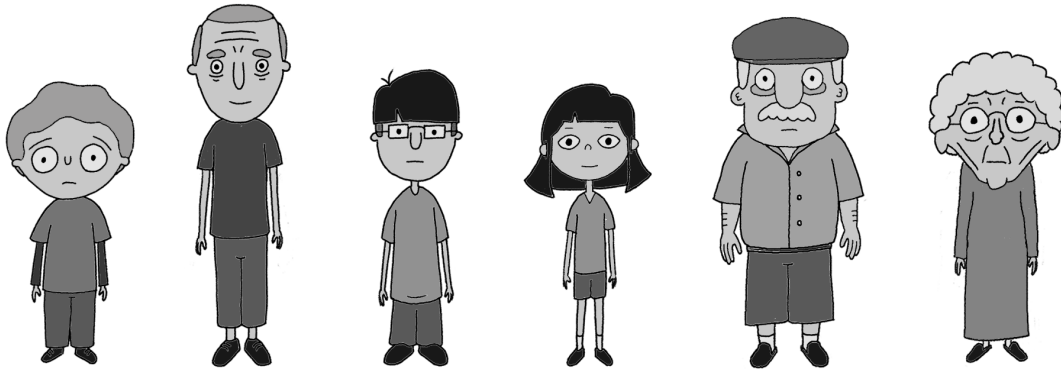
³¹ *ibidem*

³² Masaaki Yuasa, *Mind Game*

aspects qu'il prend n'est réellement le sien. Il a généralement une voix que l'on peut rattacher à son physique actuel, ce qui est logique. Nishi recherche un ensemble cohérent. Mais il ne manque pas moins d'imagination pour autant : l'aspect de Dieu prend des formes infiniment diverses et originales (un canard à chapeau, un homme à tête d'aquarium, un hippopotame gonflable pour citer quelques exemples). Cette instabilité est marquée plus encore par l'absence ponctuelle de synchronisation labiale sur le personnage : on sent qu'on est face à un artifice. La voix joue cette instabilité avec l'image, elle change sans cesse, d'une voix d'homme à une voix de femme, puis une voix très exagérée de cartoon, une voix pitchée, une voix double etc. Le personnage, complètement absurde, réussit l'exploit étrange de dégager à la fois un côté comique et une grande puissance. Cette dernière finit par prendre le dessus : alors que Nishi tente de fuir, Dieu lui court après et devient plus sérieux pour expliquer que fuir est impossible. Sa forme se stabilise alors : le corps d'un léopard avec une voix profonde et imposante.

Ce n'est donc pas parce que le cinéma d'animation peut s'écarter plus facilement des standards de la voix humaine qu'il ne peut pas jouer avec, pour créer des effets de mise en scène divers et surprenants. J'ai parlé plus tôt de l'existence de normes de voix dans la fiction et je pense que l'on peut également parler de normes de voix spécifiques à l'animation. J'en aborde d'ailleurs au début de cette partie : un personnage cartoonesque aux traits exagérés va généralement porter une voix à timbre particulier accompagnée d'une prosodie exagérée, et à l'inverse un personnage humain aux traits réalistes, va plus souvent avoir une voix proche de celles de notre quotidien. Le cinéma d'animation peut donc également jouer avec ses propres normes. Aussi, comme dans les exemples d'*Anomalisa* et de *Mind Game*, ce médium peut instaurer ses propres règles concernant l'utilisation de la voix dans le film et finir par les enfreindre pour créer un effet de mise en scène.

4 - Les voix dans mon projet de fin d'études



Mon projet de fin d'études est un court-métrage d'animation donc voici le synopsis :

Max, un jeune garçon de 10 ans, emménage dans un immeuble de banlieue-dortoir pour vivre avec son père qui y habite depuis peu. Ce dernier décide de lui préparer des pancakes pour son premier petit déjeuner dans l'appartement. Problème : il ne trouve pas de levure pour compléter sa pâte. Pour aider son père, Max part donc à la recherche d'un voisin qui pourrait leur donner un sachet. Mais il va rapidement réaliser que l'immeuble qu'il vient de rejoindre est loin d'être commode.

Dans le cadre de mon film, j'ai rapidement souhaité interpréter moi-même des personnages à la voix. J'ai d'ailleurs même considéré l'idée de tous les interpréter. J'avais déjà tenté l'expérience dans le petit court-métrage d'animation *Pipi* que nous avons réalisé dans le cadre des ateliers inter-écoles avec Arthur Chopin. Comme dans le cas de Justin Roiland, j'y interprète deux personnages avec un grand écart d'âge : un jeune homme et un vieux monsieur. Cela m'a permis d'imiter deux voix bien distinctes.

La création de l'animatique (première étape de montage du film avec des dessins et des voix provisoires) m'a permis de tester différentes voix et m'a rapidement convaincu qu'il valait vraiment mieux que je fasse appel à des comédiens. Ce, notamment, à cause du nombre de personnages et à mon incapacité à me trouver des voix satisfaisantes pour

les enfants. Le seul personnage que j'interprète dans mon film est finalement la propriétaire de l'immeuble, une grand-mère folle. Cela peut sembler contre-instinctif car il y a deux rôles d'hommes adultes dans le film, mais c'est le personnage que j'avais la sensation de mieux interpréter. Ce choix a également pour moi un intérêt narratif, je souhaitais que la propriétaire, qui est la menace centrale du film et le dernier personnage à apparaître, se démarque des autres de part son étrangeté. Ainsi tous les personnages du film ont une voix qui correspond à leur âge et à leur sexe à part la propriétaire. L'objectif étant de quitter les normes de la voix que j'ai instaurées dans mon film lors de son apparition pour créer une rupture avec les autres personnages et la mettre sur un autre plan. Comme pour les exemples cités au-dessus, dans les parties sur l'altération et l'emprunt d'une voix, l'idée est de retirer au personnage son humanité et donc d'en faire quelqu'un dont on pourrait se méfier. Je pense que l'idée de rupture apporte cela : si j'avais inscrit comme standard dans le film que tous les personnages parlent avec ma voix, celui de la propriétaire se serait inscrit dans une continuité logique. *Anomalisa* se base sur une rupture similaire. Tous les gens qui entourent Michael ont la même voix. C'est une norme que le spectateur assimile au début du film. La rupture est ressentie lorsque l'on quitte cette norme avec l'apparition de la voix de femme.

Si j'ai pu me permettre de jouer une grand-mère dans mon film, c'est notamment car mes personnages sont très loin d'être réalistes. C'est une fois de plus cette distance avec notre réalité qui m'a permis de m'affranchir des standards de la voix. Même si j'imité une voix de grand-mère, on entend facilement que ce n'en est pas une, comme on entend clairement que les enfants sont joués par des adultes dans *South Park*.

Pour le reste des personnages, je n'ai pas souhaité m'éloigner autant de notre réalité au son qu'à l'image. J'ai décidé de faire intervenir des comédiens dont l'âge et le sexe correspondaient aux personnages et de les faire jouer avec leur voix naturelle, sans tenter de modifier leur timbre. Cela était d'ailleurs déjà précisé sur les annonces de casting et j'ai été surpris de recevoir tout de même beaucoup d'enregistrements de gens qui altéraient leur timbre et exagéraient nettement leur prosodie, tentant sûrement de correspondre à une voix qu'ils pensaient adaptée au dessin animé.

Je tenais particulièrement à avoir des acteurs enfants. Il est assez courant en animation que ce soit des femmes adultes qui interprètent les enfants, comme, en France, Brigitte Lecordier qui double beaucoup de petits garçons. Je trouve ce genre de prestation vraiment remarquable, et je pense qu'il y a plein de rôles pour lesquels c'est le meilleur choix, aussi bien artistiquement, que sur le plan pratique (durée de travail, facilité de direction). Mais, j'avais pour ma part vraiment envie que les voix d'enfants soient authentiques car l'enfance est pour moi le sujet central de mon projet depuis l'écriture.

Le dessin des personnages était l'une des premières étapes de la création du film. C'est à partir de ces dessins et du texte que j'ai ensuite recherché des voix qui correspondaient selon moi aux physiques et à la personnalité des protagonistes. Je pense que leur apparence a d'ailleurs eu une grande influence sur l'idée que je me faisais des voix du film, puis, sur les choix aux castings. L'idée m'était venue de faire passer ces derniers à l'aveugle pour ne pas être influencé par le physique des comédiens. J'ai finalement renoncé à cette idée par peur de déstabiliser les acteurs (en particulier les enfants) mais je pense quand même qu'elle est intéressante. Une fois les présentations faites et les explications données, je fermais d'ailleurs les yeux ou regardais vers le bas pendant les essais pour essayer de me concentrer uniquement sur la voix.

Nous avons ensuite enregistré les acteurs en tentant le plus possible de leur faire jouer les séquences ensemble pour avoir de vrais dialogues qui s'installent. Cela était très intéressant pour trouver le ton des discussions et avoir une écoute d'ensemble de ces dernières. Nous avons ensuite repris les dialogues par petits segments pour que je puisse donner des indications plus précises et pour avoir plus de choix au montage des voix.

Comme pour les dessins avant ou après le casting, ce médium laisse le choix des enregistrements avant ou après l'animation. J'ai pour ma part décidé de commencer par enregistrer les voix. Cela pour deux raisons. Premièrement, j'aimais l'idée que les acteurs puissent se détacher du texte et proposer des variantes. Deuxièmement, ne travaillant pas exclusivement avec des acteurs qui s'étaient déjà essayés au doublage, je ne voulais pas que l'exercice de la bande rythmo (outil de défilement de texte sous l'écran permettant de jouer au bon rythme) complique leur prestation.

La bande rythmo a par contre ses propres avantages. Elle permet aux comédiens de pouvoir jouer face à l'animation terminée, les aidant ainsi à pouvoir ajuster leur jeu à l'image et voir directement l'interaction entre les deux. Le réalisateur peut alors travailler en direct avec les comédiens pour obtenir le résultat qu'il souhaite pour son film en pouvant avoir une vue d'ensemble sur le projet.

CONCLUSION

Nos dialogues au quotidien sont régis par des standards de conversations et nos voix par des normes physiques et de prosodie. Toutes ces normes, qui facilitent nos interactions, influencent la manière dont la voix est utilisée au cinéma. En effet, suite à l'arrivée du parlant, le cinéma est rapidement devenu vococentriste. Pour parler de cette révolution on utilise le mot "parlant" et non "sonore". La voix est souvent au cinéma, un vecteur d'informations et d'émotions (comme au quotidien). Pour permettre cela, l'ensemble de la chaîne du son s'attelle, la plupart du temps, à faire en sorte que la voix soit intelligible mais également crédible pour que le spectateur croie en son existence dans l'espace du film. Ces deux critères peuvent être respectés de très nombreuses manières et laissent une très grande marge artistique aux métiers du son concernant le traitement de la voix. Le choix du ou des points d'écoute est d'ailleurs un outil de mise en scène qui peut fortement influencer le ressenti du spectateur, en grande partie de manière inconsciente. Un autre outil avec lequel on peut repenser notre approche de la voix est la prosodie. On pourrait se dire qu'un acteur joue mal s'il ne parle pas comme dans la vraie vie, autrement dit, si sa prosodie ne correspond pas à celle que l'on entend au quotidien. Mais ce n'est pas forcément le cas. Un cinéaste n'est pas forcément obligé de penser que la représentation du monde va de soi. Il n'est pas obligé de considérer que le dialogue se doit d'instaurer un climat réaliste. Par exemple, Robert Bresson instaurait à ses modèles une manière de parler propre à ses films. Cette dernière, bien que éloignée de nos habitudes, était pour lui la plus juste. Elle avait pour but d'être la plus automatique possible car, d'après lui «C'est de l'automatisme, qui compose les trois quarts de notre existence, que sort le vrai, et non du pensé et du raisonné.»³³. Il est également possible d'isoler la prosodie et d'en faire un élément de narration plus important encore que les mots prononcés. Les petits fragments de voix souvent presque

³³ Pierre Ajame, *Le cinéma selon Bresson*

inaudibles de Monsieur Hulot, le personnage de Jacques Tati, suivent ce procédé. Ces derniers font partie intégrante de la silhouette du personnage et aide à le caractériser.

Il est intéressant de noter que même si des réalisateurs altèrent sciemment la prosodie, on peut considérer que l'exercice de création d'un film le fait aussi souvent automatiquement. Les procédés d'écriture, d'apprentissage et de jeu des dialogues font que la prosodie des acteurs est souvent plus fluide, mieux construite, que la nôtre.

Un autre paramètre qui peut être modifié pour le bien de la mise en scène est la nature de la voix elle-même : sa hauteur, son intensité et son timbre. Si on prête une voix humaine à un personnage qui ne l'est pas, il est courant que sa voix soit modifiée pour correspondre à sa taille, à sa morphologie, à sa nature. Mais il est intéressant de remarquer que lorsque qu'il s'agit d'un personnage auquel le metteur en scène souhaite donner une part d'humanité, une question d'équilibre apparaît. Un équilibre subtil entre la crédibilité du rapport corps/voix du personnage et son besoin de proximité avec la voix humaine. Il semble qu'une des meilleures réponses à cette question soit de ne pas uniquement compter sur la post-production pour trouver la voix du personnage. Il faut la chercher le plus tôt possible et si possible intégrer l'acteur au processus de recherche. Plus ce dernier sera apte à incarner son rôle par la voix, moins il sera difficile d'adapter cette dernière à la créature qu'il interprète. Il est intéressant de noter que, dans de nombreux films, la mise en scène ne modifie simplement pas la voix de la créature qu'elle souhaite humaniser, malgré un rapport corps/voix pouvant paraître surprenant. Mais prêter une voix humaine à autre chose qu'un humain peut avoir un effet complètement inverse et entraîner une sensation de malaise ou même de peur chez un spectateur.

Les modificateurs de voix peuvent aussi être prêtés à des personnages humanoïdes. Pour les créatures évoquées au-dessus, une voix humaine, même altérée, a tendance à donner de l'humanité. C'est complètement l'inverse pour les protagonistes à l'aspect humain. S'ils quittent les standards de la voix, ils perdent instantanément toute humanité. Dans le cas où la raison de ce phénomène n'est pas encore justifiée, on peut rapidement supposer qu'ils ont sûrement d'autres capacités surnaturelles.

Mais la mise en scène n'est pas obligée de compter sur l'altération de la voix pour créer des rapports marginaux entre le corps et les voix. Elle peut par exemple utiliser la désynchronisation. Élément qui a longtemps été facile à accepter pour les spectateurs

mais qui est aujourd'hui un critère important de crédibilité. Marguerite Duras dans *India Song* assume une désynchronisation totale entre le film des voix et le film des images. Les deux se déroulant à des rythmes différents. Cela donne un aspect très impalpable au film qui nous permet de sentir que ses événements sont de l'ordre du souvenir.

Il est également possible de donner une voix provenant d'un corps à un autre corps. Lorsque ce transfert contredit les normes vocales (une voix trop grave pour une personne trop jeune par exemple), on retrouve un effet de déshumanisation.

Le réalisateur a donc de nombreux outils à sa disposition pour servir sa narration en quittant les standards de la voix. Il peut se réapproprier la prosodie, altérer les voix et jouer avec le rapport corps/voix. Ces trois méthodes, qui sont souvent des partis pris marqués dans le cinéma en prise de vue réelle, sont très courantes et parfois complètement normalisées en animation. Cela est notamment facilité par le fait qu'il n'y ait aucun rapport corps/voix forcé en animation et que ce médium n'est pas une captation de notre monde mais une reproduction plus ou moins proche de ce dernier. Ainsi, il semble que plus l'univers animé est éloigné du nôtre, plus il est facile de s'affranchir des normes de la voix humaine. Il n'y a, toutefois, aucune nécessité de s'éloigner autant de notre réalité par l'image et par le son en animation. Très courants sont les exemples de film avec un univers différent du nôtre dans lesquels des personnages très éloignés de nous (aussi bien visuellement que par leur nature) sont interprétés par des acteurs qui jouent avec une voix naturelle. La voix humaine étant un fort vecteur d'émotions qui facilite le développement d'empathie pour un personnage, quelle que soit sa forme. L'animation peut donc s'affranchir plus facilement de la voix humaine. Malgré tout, comme le cinéma en prise de vue réelle, ce médium a lui aussi des normes de voix qui se sont construites avec le temps. Cela lui permet de jouer avec ses propres standards pour mettre en place des effets de mise en scène se basant sur la voix humaine. Il est aussi possible pour un film d'animation d'établir ses propres règles concernant la voix et de les enfreindre pour servir la narration. C'est ce que je tente de faire dans mon projet de fin d'études : tous les personnages du film ont une voix qui correspond à leur âge et leur sexe. Je brise cette règle à l'apparition du dernier personnage, une femme âgée qui porte ma voix. L'objectif étant de déshumaniser ce personnage et de l'isoler du reste des protagonistes par l'étrangeté de son rapport corps/voix, le plaçant alors comme une anomalie aux yeux des spectateurs.

BIBLIOGRAPHIE

Ouvrages

VASSÉ Claire, *Le dialogue : du texte écrit à la voix mise en scène*, Cahiers du Cinéma, Paris, 2003.

BELLER Grégory, *La musicalité de la voix parlée*, mémoire de maîtrise en musique, Université Paris 8, Paris, 2005.

DURAS Marguerite, *India Song*, Gallimard, Paris, 1973.

CHION Michel, *La Voix au cinéma*, Cahiers du Cinéma / Éditions de l'Étoile, Paris, 1982.

CORNUT Guy, *La voix*, Presses Universitaires de France (PUF), Paris, 2019.

BRESSON Robert, *Notes sur le cinématographe*, Gallimard, Paris, 1975.

CHION Michel, *David Lynch*, Cahiers du Cinéma / Éditions de l'Étoile, Paris, 2007.

FELLINI Federico, *Fellini par Fellini*, entretiens avec Giovanni Grazzini, traduit de l'italien par Nino Frank, Calmann-Lévy, Paris, 1984.

Articles

GRICE Herbert Paul, *Logic and Conversation*, dans P. Cole et J.-L. Morgan (dir.), *Syntax and Semantics*, vol. 3 : *Speech Acts*, Academic Press, New York, 1975, p. 41–58.

SCHERER Klaus R., *Expression of Emotion in Voice and Music*, *Journal of Voice*, vol. 9, n° 3, 1995, p. 235–248.

DELATTRE Pierre, *Les Dix Intonations de base du français*, *The French Review*, vol. 40, n° 1, octobre 1966, p. 1–14.

AJAME Pierre, *Le cinéma selon Bresson*, *Les Nouvelles Littéraires*, 26 mai 1966, p. 1–13.

Conférences

CHARVET Alexandre, *La voix et ses métamorphoses dans « Les Métamorphoses » d'Ovide*, conférence, 2004.

THÉLOT Jérôme, *Voix et sacrifice dans le cinématographe selon Robert Bresson*, intervention au colloque BnF / ARM, Paris, 5 décembre 2015.

Vidéos

Youtube : SoundWorks Collection - The Sound of The Hobbit
<https://www.youtube.com/watch?v=KcOz5S3fKBE>

Youtube : Making-of Linda veut du poulet : la prise de son
<https://www.youtube.com/watch?v=nftzsvmT9Zs>

Filmographie

Films

GUADAGNINO Luca, *Queer*, 2025.

BRESSON Robert, *Au Hasard Balthazar*, 1966.

BRESSON Robert, *Un condamné à mort s'est échappé*, 1966.

LYNCH David, *EraserHead*, 1977.

TATI Jacques, *Playtime*, 1967.

GLAZER Jonathan, *Under the skin*, 2013

COLUMBUS Chris, *Harry Potter et la chambre des secrets*, 2002.

HILL Tim, *Alvin et les Chipmunks*, 2007.

JACKSON Peter, *Le Seigneur des anneaux : Les Deux Tours*, 2002.

WHEDON Joss, *Avengers*, 2012.

JACKSON Peter, *Le Hobbit : La désolation de Smaug*, 2013.

CAILLEY Thomas, *Le Règne Animal*, 2023.

SCHAFFNER Franklin, *La Planète des Singes*, 1968.

SPIELBERG Steven, *Bon Gros Géant*, 2016.

DEL TORO Guillermo, *Hellboy*, 2004.

CRONENBERG David, *Le Festin Nu*, 1991

LUCAS George, *Stars Wars*, 1977

NOLAN Christopher, *Interstellar*, 2014

FAVREAU Jon, *Iron Man*, 2008

JONZE Spike, *Her*, 2013

KUBRICK Stanley, *2001, l'odyssée de l'espace*, 1968

VAUGHN Matthew, *X-MEN le commencement*, 2011

VILLENEUVE Denis, *Dune*, 2021

DURAS Marguerite, *India Song*, 1975

GIMÉNEZ PEÑA Juanjo, *En Décalage*, 2021

HITCHCOCK Alfred, *Blackmail*, 1929

KELLY Gene et DONEN Stanley, *Singin' In The Rain*, 1953

FRIEDKIN William, *L'exorciste*, 1973

FREAND JONES Nick *The Fear Of God - 25 Years Of The Exorcist*, 1998

MINKOFF Rob et ALLERS Roger, *Le Roi Lion*, 1994

MALTA Chiara et LAUDENBACH Sébastien, *Linda veut du poulet !*, 2023

LALOUX René, *La Planète Sauvage*, 1973

KAUFMAN Charlie et JOHNSON Duke, *Anomalisa*, 2015

YUASA Masaaki, *Mind Game*, 2004.

Courts-métrages

ZHYZHKO Olaa, *Just Drifting*, 2025.

FAVÉ Caroline, *Toxic Underground*, 2025.

LYNCH David, *The Grandmother*, 1969.

Séries

PARKER Trey et STONE Matt, *South Park*, 1997

ROILAND Justin et HARMON Dan, *Rick et Morty*, 2013

BOB-WAKSBERG Raphael, *Bojack Horseman*, 2014

Vidéos

FRAYSSINET Roman, *Conversation*, 2002.

https://youtu.be/zWHD_FBJAyU?si=PQ-DNP0hUHf8yRaK

FRAYSSINET Roman, *Ô dedans*, 2025.

<https://youtu.be/AmBJ90ExhFI?si=z6gqbeODk-tjApeG>

WURTZ Bill, *2 friends on the beach*, 2023

<https://youtu.be/Eklb2Qp7ItY?si=3b7XvHlM2fsfP8nT>

FIRTH David, *Cream*, 2018

<https://youtu.be/0UgiJPnwtQU?si=IL9v07LiCki1dgwj>

FIRTH David, *Gice*, 2024

https://youtu.be/fxW8zkSuvOQ?si=MIJdM_2kxX7NlZmn

Remerciements

Jacques Descomps, Térésa Faucon, Tania Press, Valérie Deloof, Jean-Pierre Laforce, Théo Serror, Pierre Sans, Nathalie Pichon, Armel Le Meur, María Mónica Martínez, Malo Geffraye Meriadec, Juliette Ragot, Pierre Locquet, Younes Nefoussi.