HAL
open science

Prédiction du déficit d'intelligibilité chez des patients traités pour un cancer de la cavité buccale ou de l'oropharynx

Laura Monestier

To cite this version:

Laura Monestier. Prédiction du déficit d'intelligibilité chez des patients traités pour un cancer de la cavité buccale ou de l'oropharynx. Linguistique. 2020. dumas-03220389

HAL Id: dumas-03220389

https://dumas.ccsd.cnrs.fr/dumas-03220389

Submitted on 7 May 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Prédiction du déficit d'intelligibilité chez des patients traités pour un cancer de la cavité buccale ou de l'oropharynx

Laura MONESTIER
Sous la direction de Muriel LALAIN
et du Professeur Nicolas FAKHRY

Master 2 Sciences du Langage

Aix-Marseille Université, CNRS, Laboratoire Parole et Langage, Aix-en-Provence, France

Septembre 2020

REMERCIEMENTS

Je tenais à remercier les personnes qui m'ont accompagnée tout au long de cette année, que ce soit au premier semestre pour m'épauler dans ces nombreuses disciplines nouvelles pour moi, ou au second dans la réalisation et l'écriture de ce mémoire.

Tout d'abord, un grand merci à Muriel Lalain. Tu as été d'une patience folle, toujours prête à m'expliquer les secrets de la linguistique avec les mots justes. Je suis ravie d'avoir pu travailler avec toi, merci de ta gentillesse et de ta pédagogie. Et surtout, merci car grâce à toi mes fichiers Word ne feront plus 130 Mo.

Merci au Professeur Fakhry, pour m'avoir accordé votre confiance et pour m'accompagner tout au long de mon internat. C'est une grande chance de pouvoir apprendre à vos côtés, et d'être encadrée par vous dans tous mes projets. Je m'arrêterai là, car il faut en garder un peu pour la thèse !

Merci à Alain Ghio, pour ta bonne humeur et ton implication. Grâce à toi les statistiques n'auront pas eu raison de moi. Je garderai en mémoire ce débat passionné entre toi et Muriel sur les consonnes palatales et post-alvéolaires !

Merci à Marie Rebourg, pour avoir été toujours disponible lorsque j'avais une question (ou besoin d'inspiration pour une idée de projet fictif en phonologie de laboratoire…). J'ai eu la chance de pouvoir m'appuyer sur tes travaux de mémoire de master 1 et de master 2. Je te souhaite plein de réussite dans ta thèse et dans ta vie personnelle.

Un grand merci à Christelle Portes, pour votre confiance et votre gentillesse tout au long de cette année.

Merci au Professeur Giovanni, au Docteur Robert ainsi qu'à toute l'équipe POP, pour leur accueil au sein de l'équipe, ainsi que leurs conseils et leurs remarques avisées concernant ce mémoire.

Et enfin un grand merci au jury d'avoir accepté d'évaluer ce travail.

Table des matières

INTRODUCTION *5*

I. LE TRACTUS VOCAL AU CARREFOUR DE L'ORL ET DE LA LINGUISITIQUE *6*

I.1. Voies aéro-digestives supérieures : anatomie et fonctions **6**

I.2. Cancers des voies aérodigestives supérieures **10**

I.3. Intelligibilité de la parole **12**

I.4. Projets C2SI et RUGBI **14**

II. PROBLEMATIQUES *16*

III. MATERIELS ET METHODES *17*

III.1. Corpus **17**

III.2. Un test de jugement perceptif de l'intelligibilité : la tâche de DAP **18**

III.3. Évaluer l'intelligibilité avec la tâche de DAP **20**

III.3.1. Phase de production 20

III.3.2. Phase de perception 20

III.3.3. Calcul du score d'intelligibilité 21

III.4. Traitement des données **23**

III.5. Analyse statistique **26**

IV. RESULTATS *27*

IV.1. Altération globale de l'intelligibilité en fonction de la taille et de la région d'une tumeur : scores PPD globaux **27**

IV.2. Altération précise des segments de parole : scores PPD locaux **28**

V. DISCUSSION *36*

REFERENCES *43*

INTRODUCTION

Les cancers des voies aéro-digestives supérieures (VADS) sont un vrai problème de santé publique. Ils représentent la 5ème cause de cancer en France, et sont dus majoritairement au tabac et à l'alcool. La moitié sont des cancers de la cavité buccale et de l'oropharynx. Depuis quelques années, nous assistons à des progrès considérables en termes de traitements, notamment avec des chirurgies plus conservatrices et plus ciblées, et des traitements médicaux (chimiothérapie et radiothérapie) plus efficaces. L'espérance de vie des patients a ainsi augmenté. Mais il persiste un problème majeur : les séquelles post-thérapeutiques, et notamment celles sur la parole. C'est ainsi à l'intersection de la cancérologie ORL et de la communication parlée que se situe notre travail. Il s'agira ici d'aborder la question de l'évaluation de l'intelligibilité de la parole en mêlant nos connaissances cliniques à nos connaissances linguistiques plus récemment acquises. En effet l'intelligibilité de la parole est altérée d'une part par la tumeur elle-même mais aussi par les traitements du cancer. Cependant, il existe encore des zones d'ombres d'une part sur la caractérisation du lien entre une région anatomique et un déficit fonctionnel donné et d'autre part sur la qualité des mesures permettant d'évaluer l'intelligibilité. A travers ce mémoire de Master, nous essaierons de contribuer à la réflexion sur la prise en charge des séquelles communicationnelles qu'impliquent les traitements pour un cancer de la cavité buccale et de l'oropharynx. En nous appuyant sur les données de deux projets financés (les projets C2SI et RUGBI), nous nous intéresserons ici à une tâche d'évaluation de l'intelligibilité récemment développée. Notre objectif sera de contribuer à la validation de cette tâche en examinant (i) sa capacité à prédire le degré d'altération de l'intelligibilité en fonction de la taille et de la localisation d'une tumeur et (ii) sa capacité à prédire quels sons de parole sont les plus susceptibles d'être altérés en fonction de ces mêmes facteurs cliniques.

Nous donnerons dans un premier temps les éléments théoriques qui nous ont permis de développer notre problématique. Nous décrirons ensuite les matériels et méthodes utilisés. Enfin, nous présenterons et discuterons les résultats obtenus dans cette étude.

I. LE TRACTUS VOCAL AU CARREFOUR DE L'ORL ET DE LA LINGUISITIQUE

Le tractus vocal abrite plusieurs fonctions, biologiques vitales et communicationnelles, qui intéressent respectivement cliniciens et linguistes. Dans le cadre de ce travail, le tractus vocal sera le siège, du point de vue de l'ORL, de tumeurs cancéreuses dont les traitements, par les modifications anatomiques qu'ils induisent, viendront impacter la parole produite en son sein. Nous verrons dans cette première partie, comment les structures anatomiques sont appréciées, d'un point de vue ORL et linguistique, ce qui nous permettra de situer la problématique de l'évaluation de l'intelligibilité chez des patients traités pour un cancer de la cavité buccale ou de l'oropharynx. Nous rappellerons ci-après ce que sont la cavité buccale et l'oropharynx et leur implication dans la parole, les points importants concernant les cancers de ces deux zones anatomiques et leurs conséquences sur l'intelligibilité, nécessitant donc l'évaluation de cette dernière.

I.1. Voies aéro-digestives supérieures : anatomie et fonctions

Les voies aérodigestives supérieures (VADS) représentent la partie haute de l'appareil respiratoire et du système digestif. Elles comprennent les fosses nasales, la cavité buccale, le rhinopharynx, l'oropharynx, l'hypopharynx, le larynx, la partie supérieure de la trachée et la partie supérieure de l'œsophage. Les VADS ont donc deux fonctions biologiques majeures : la première est respiratoire en faisant circuler l'air jusqu'à la trachée puis aux poumons, après l'avoir réchauffé et filtré, et la deuxième est digestive en acheminant le bol alimentaire jusqu'à l'œsophage puis l'estomac (mastication puis déglutition). Mais ce ne sont pas ses uniques fonctions. En effet les VADS participent également à l'odorat, à la gustation et ce qui va nous intéresser plus particulièrement dans ce mémoire, à la phonation.

Nous nous intéresserons donc tout particulièrement à la cavité buccale et à l'oropharynx, car bien que d'autres régions des VADS soient impliquées dans la phonation (notamment le larynx avec les cordes vocales), la parole à proprement parler dépend de ces deux zones. La cavité buccale comprend les lèvres (excepté leur versant cutané), la langue mobile, le plancher buccal, le palais osseux, les faces internes de joue et la région rétromolaire. L'oropharynx est la région limitée en haut par le voile du palais, en bas par un plan horizontal passant par le bord supérieur de l'épiglotte et l'os hyoïde, en avant par l'isthme du gosier et le V lingual et en arrière par la paroi pharyngée postérieure. Il comprend ainsi la base de langue, les amygdales

palatines, le voile du palais et la paroi pharyngée postérieure (Fig.1). On parle parfois de carrefour aéro-digestif, car c'est au niveau de l'oropharynx que les voies digestive et respiratoire se rencontrent.

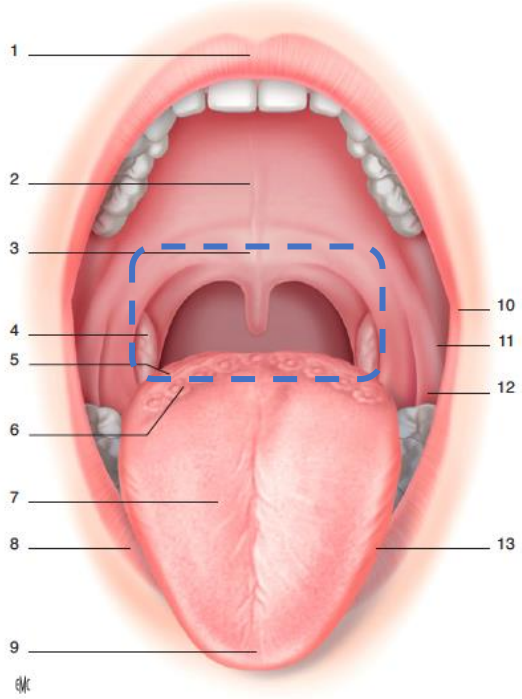

Figure 1 : Cavité buccale et oropharynx (bouche ouverte, langue tirée). 1. Lèvre supérieure ; 2. palais dur ; 3. voile du palais ; 4. amygdale ; 5. base de langue ; 6. V lingual ; 7. dos de la langue ; 8. lèvre inférieure ; 9. pointe de la langue ; 10. Commissure labiale ; 11. face interne de joue ; 12. trigone rétromolaire ; 13. bord latéral gauche de la langue ; pointillés bleus : oropharynx (Encyclopédie médico-chirurgicale d'Oto-Rhino-Laryngologie, 2018 : volume 13, n°4)

Au cours de la parole, c'est tout un système articulatoire qui va se mettre en place afin de produire la parole, dont le siège est précisément le système supralaryngé qui comprend le pharynx et la cavité buccale.

La production de la parole est un phénomène complexe, impliquant plus de 100 muscles différents et plusieurs organes : les poumons, le larynx et le conduit vocal notamment. Ces organes s'agencent en plusieurs étages, que l'on pourrait diviser en trois fonctions à savoir la soufflerie, la phonation et l'articulation. En effet, les poumons, responsables de l'oxygénation du sang dans l'organisme, vont expulser une certaine quantité d'air lors de l'expiration, air qui va circuler via la trachée jusqu'au larynx. Au cours de la phonation, le larynx, et plus précisément les deux cordes vocales, ouvertes lors de la respiration, vont pouvoir se fermer et vibrer au passage de l'air expiré. Cela va générer un son, qu'on caractérisera de voisé (ou sonore). Lorsqu'un son n'implique pas la vibration des cordes vocales, il est dit non voisé (ou sourd). Le son produit grâce au souffle des poumons et plus ou moins aux mouvements des cordes vocales va ensuite circuler dans le conduit vocal afin d'y être modulé, grâce à deux types d'éléments : les résonateurs et les articulateurs. Les résonateurs, que sont la cavité buccale et la

cavité nasale, vont jouer comme des caisses de résonance créant respectivement des sons oraux ou nasaux. Les articulateurs (mobiles avec la langue, la mandibule, le voile du palais et les lèvres, et immobiles avec les dents et le palais dur) vont encore affiner ce son à l'aide de mouvements plus ou moins fins. Ces mouvements vont venir modifier la forme et le volume des résonateurs, donnant ainsi un timbre propre à chacun des sons de parole. La langue présente par exemple des mouvements antéro-postérieurs et verticaux permettant de jouer sur différents traits des sons comme le lieu d'articulation, le mode articulatoire, l'aperture ou la position antéro-postérieure d'une voyelle. La mandibule à l'aide d'un mouvement d'ouverture est également responsable de l'aperture d'une voyelle. Enfin le voile du palais, lorsqu'il s'ouvre (ou se ferme) permettra le passage de l'air dans la cavité nasale (ou non) créant ainsi un son nasal (ou oral).

Ainsi en fonction de l'utilisation de chacun de ces organes, on peut attribuer à chaque son produit différents traits articulatoires, constituant comme une carte d'identité du phonème. En français, on retrouve deux types de sons : les sons consonantiques et les sons vocaliques. Les consonnes sont définies selon quatre traits :

- le lieu d'articulation : il renseigne l'endroit où se rencontrent les articulateurs dans le conduit vocal (par exemple : bilabial, labiodental)
- le mode d'articulation : il est relatif au degré de constriction du conduit vocal, qui sera respectivement total ou partiel pour une consonne occlusive et une fricative. Dans les deux cas le flux d'air est donc turbulent (à l'inverse des voyelles pour lesquelles le flux d'air est laminaire)
- la résonance : nasale ou orale, dépend directement de la position du voile du palais
- le voisement : sonore (les cordes vocales vibrent) ou sourde (ne vibrent pas)

Les voyelles sont définies par les quatre traits suivants :

- l'aperture : langue et mandibule plus ou moins basses
- la position de la langue sur l'axe antéro-postérieur
- l'arrondissement des lèvres
- le caractère oral/nasal

Par défaut, les voyelles sont toutes voisées.

La figure 2 illustre cette notion de traits, sur laquelle nous prendrons appui au cours de nos analyses. A gauche, on retrouve une coupe sagittale du conduit vocal lors de l'articulation de /n/ qui est une consonne occlusive apico-alvéolaire sonore nasale. A droite, on retrouve la coupe sagittale lors de l'articulation de /d/, qui est une consonne occlusive apico-alvéolaire sonore orale. Ces deux consonnes ne diffèrent donc que d'un trait, celui de la résonance. Pour le /d/ le voile du palais est en effet fermé, alors qu'il est ouvert pour le /n/, permettant ainsi de passer

d'un seul résonateur (nasal) à deux (nasal et oral). Le /n/ est ainsi un phonème nasal alors que le /d/ est un phonème oral. Cette notion de trait d'écart est extrêmement importante lorsque l'on parle de phonème, ce dernier étant défini comme la plus petite unité linguistique permettant d'opposer du sens. Certains phonèmes ont un seul trait d'écart entre eux, un seul trait d'écart peut donc changer tout le sens d'un mot.

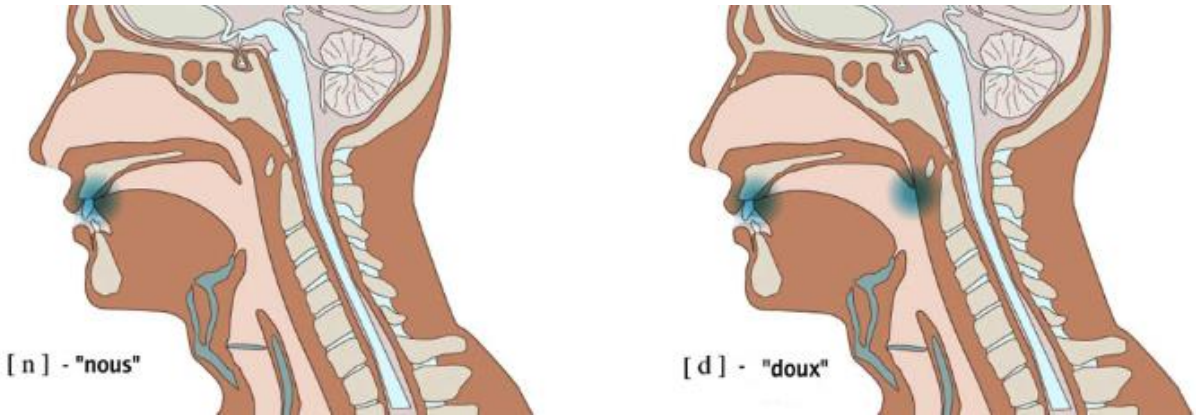

Figure 2 : Coupes sagittales du tractus vocal lors de l'articulation des phonèmes /n/ (à gauche) et /d/ (à droite). Les zones bleues indiquent les zones d'occlusion. Pour le /n/ la langue va entrer en contact avec les dents (apico-alvéolaire) mais le voile est ouvert, alors que pour le /d/ on observe le contact apico-alvéolaire et le voile du palais en fermeture, venant bloquer le passage de l'air dans la cavité nasale.

Dans ce travail portant sur l'intelligibilité et l'articulation des sons, nous mettrons de côté les poumons et le larynx, nous intéressant tout particulièrement aux résonateurs et articulateurs. En effet on imagine facilement qu'une altération d'un de ces éléments, que l'on rencontre notamment au cours d'un cancer de la cavité buccale ou de l'oropharynx, notre contexte clinique, engendre une modification de l'articulation des sons et donc de l'intelligibilité de la parole. Si nous reprenons l'exemple du /d/ et du /n/ ci-dessus, il est évident qu'une phrase comme « Nous validons cela » n'a rien à voir avec « Doux validons cela », ou encore « Oui, votre Honneur » et « Oui, votre odeur ». Ce qui peut nous paraître comme une infime erreur peut donc se révéler à l'échelle humaine comme terriblement handicapante sur le plan communicationnel. On peut imaginer qu'un patient atteint d'un cancer du voile du palais a probablement une mobilité altérée de celui-ci, et peut donc être très gêné pour les phonèmes impliquant la résonance nasale. La figure 3 illustre le rapport qu'il existe entre la région anatomique (cavité buccale ou oropharynx) et les phonèmes : certains seront réalisés au niveau de la cavité buccale, d'autres dans l'oropharynx. De là découlent nos intuitions : une atteinte de la cavité buccale altèrerait plutôt les phonèmes antérieurs, comme ceux qui sont réalisés au niveau labiodental ou alvéolaire par exemple. Alors qu'une atteinte de l'oropharynx dégraderait principalement les phonèmes de réalisation postérieure, comme ceux articulés au niveau vélaire ou ceux présentant une résonance nasale.

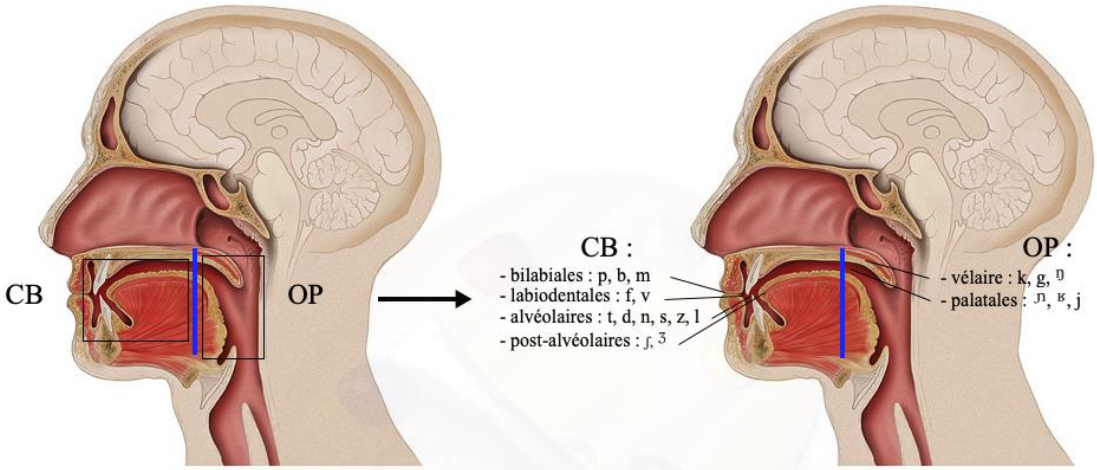

Figure 3 : Coupes sagittales du tractus vocal. A gauche, le trait bleu symbolise la délimitation entre la cavité buccale et l'oropharynx. A droite, on peut voir les lieux d'articulation des phonèmes consonantiques en fonction de la région anatomique concernée

I.2. Cancers des voies aérodigestives supérieures

Les cancers des VADS représentent la 5ème cause de cancer en France (Balaguer et al., 2019). Les cancers de la cavité buccale représentent environ un tiers de ceux-ci, et ceux de l'oropharynx environ 15-20%. Ils se déclarent généralement chez des hommes de 45-55 ans présentant un tabagisme chronique et/ou une consommation alcoolique, qui sont les deux facteurs de risque principaux. Mais le sex ratio de cette pathologie tend à s'inverser, la consommation tabagique étant moins fréquente chez les hommes et de plus en plus répandue chez les femmes. De plus, on assiste depuis quelques années à un essor des cancers de l'oropharynx non pas dus à une consommation alcoolo-tabagique mais à une infection virale à Human Papilloma Virus (HPV), que l'on retrouve notamment dans le cancer du col de l'utérus. Ces cancers HPV-induits sont de meilleur pronostic, et peuvent se déclencher chez des gens jeunes. Les symptômes amenant à consulter sont une douleur, une gêne à l'alimentation ou à la parole, des saignements, ou des signes plus généraux comme une fatigue ou une perte de poids. Parfois le patient a aperçu lui-même une lésion charnue, bourgeonnante ou creusante dans sa cavité buccale (Fig.4).

Figure 4 : Cancer de la face inférieure gauche de la langue. Aspect bourgeonnant, irrégulier de la muqueuse.

Suite à la première consultation, un bilan complet est réalisé. Il comprend des imageries médicales (scanner, parfois IRM) afin de faire le point sur la lésion, son extension aux tissus adjacents mais aussi pour rechercher des lésions à distances (métastases), des biopsies vont être réalisées, s'il le faut sous anesthésie générale, ainsi qu'un bilan nutritionnel, parfois psychologique. A l'issue de ce bilan, le dossier du patient est discuté en réunion de concertation pluridisciplinaire, où tous les acteurs de la prise en charge sont présents : ORL, radiologues, chimiothérapeutes, radiothérapeutes, … La tumeur est classée selon sa taille (T), le nombre de ganglions atteints (N) et la présence de métastases ou pas (M). En fonction de cette classification TNM, de l'état général du patient, de la localisation de la tumeur et d'autres facteurs encore, le traitement dont va bénéficier le patient va être décidé de manière collégiale.

En effet les cancers des VADS peuvent se traiter de plusieurs manières. La chirurgie par exemple, qui vise à retirer la tumeur dans son intégralité avec des marges de tissu sain autour, que l'on appelle marges de sécurité. La perte de substance qui en découle, si elle est trop importante pour être suturée directement, ou être corrigée en réalisant un lambeau (qui sera constitué de peau, de muqueuse, de muscle voire même d'os selon les besoins), que l'on prélèvera juste à côté ou à distance de l'exérèse. La chirurgie peut être associée à ou remplacée par de la radiothérapie et/ou de la chimiothérapie. Au fil des années, tous ces traitements ont vu leurs techniques évoluer, les chirurgies sont moins mutilantes, si bien que le pronostic de ces cancers s'est considérablement amélioré et les répercussions sont dorénavant moins lourdes. Mais malgré un taux de guérison en amélioration, les conséquences fonctionnelles restent malgré tout conséquentes chez ces patients, et notamment sur la déglutition et la parole (McKinstry et al., 2003). Il en résulte une altération franche de la qualité de vie. Au cours du suivi du patient qui fera suite au traitement, l'enjeu sera évidemment de guetter la moindre récidive cancéreuse, mais aussi d'évaluer cette altération de la qualité de vie, par l'évaluation

de la déglutition et de la parole, d'une part pour orienter la prise en charge thérapeutique puis évaluer son efficacité (orthophonie), mais aussi pour informer le patient de l'évolution probable de sa maladie, tout cela dans le cadre d'une prise en charge optimale.

I.3. Intelligibilité de la parole

Dans la pratique clinique, et notamment en cancérologie ORL, le clinicien va ainsi devoir évaluer le trouble de production de la parole dont souffre son patient, et cela passe par l'évaluation de l'intelligibilité. Celle-ci comporte trois dimensions : la première est en lien avec les performances du locuteur, la seconde correspond au canal de transmission entre le locuteur et l'auditeur, et enfin la dernière reflète les capacités de l'auditeur. Le versant qui nous intéresse pour ce travail est celui en lien avec le locuteur, puisque le but à terme est d'évaluer la capacité à communiquer d'un patient présentant un trouble de la parole post-traitement d'un cancer de la cavité buccale ou de l'oropharynx.

Il y a plusieurs définitions de l'intelligibilité. Certains la définissent comme le degré de précision avec lequel un message est compris par l'auditeur (Yorkston, Dowden et Beukelman, 1992). Mais cette définition peut être jugée imprécise par certains, puisque cela englobe les trois aspects différents précédemment évoqués de l'intelligibilité. Virginie Woisard est allée plus loin, en faisant bien la distinction entre intelligibilité et compréhensibilité (Rebourg, 2017). L'intelligibilité est la précision avec laquelle un auditeur récupère le signal acoustique produit par le locuteur, alors que la compréhensibilité est la capacité de l'auditeur à interpréter le sens du message oral produit par le locuteur sans tenir compte de la précision ou de la justesse phonétique ou lexicale. Rebourg et al. (2020) adoptent plutôt la définition de Keintz, Bunton et Hoit (2007), qui définissent l'intelligibilité comme « la quantité de parole comprise à partir du seul signal acoustique ». Cette dernière établit ainsi une relation étroite entre intelligibilité et informations dépendantes du signal (bottom up) liées au processus de décodage des sons. Notre travail visant à analyser la capacité communicationnelle des patients atteints d'un cancer de la cavité buccale ou de l'oropharynx, cette dernière définition nous paraît totalement adaptée.

Mais comment évalue-t-on l'intelligibilité ? C'est une question primordiale, puisque c'est sur cette évaluation que va se baser toute la prise en charge thérapeutique de la parole qui va suivre, notamment la rééducation orthophonique. Or, les cliniciens ne disposent pas d'outil d'évaluation standardisé, plusieurs tests sont utilisés mais aucun n'est satisfaisant pour l'instant. Le gold standard reste l'évaluation perceptive par un auditeur. Ceci passe par des systèmes de notation par un ou plusieurs auditeurs, sur une échelle de 0 à 10 par exemple, portant sur le discours spontané d'un locuteur, une description d'image ou de la lecture (Astesano et al, 2018),

ou encore la répétition de syllabes type « ka », « va », « da », etc (Borrgreven et al, 2005). Il peut s'agir aussi du calcul du pourcentage d'unités de paroles reconnus par l'auditeur, mais d'autres cliniciens utilisent également des épreuves extraites de batteries de tests. On peut citer, de manière non exhaustive, le test de diagnostic par paires minimales (Peckels et Rossi, 1973), qui a le mérite d'être bien contrôlé au niveau phonologique mais très long et ne testant que les consonnes en position initiale. Le test phonétique d'intelligibilité de la Batterie d'Évaluation Clinique de la Dysarthrie (BECD), lui, est fréquemment utilisé mais comporte 50 mots, mono ou bisyllabiques. Une nouvelle version a vu le jour, et est constituée cette fois-ci de mots plurisyllabiques donc paraît déjà plus adaptée (Blanc et Giusti, 2014). Aussi nombreux soient ces tests, aucun n'est validé et n'obtient de consensus pour une application clinique satisfaisante. En effet, tous se basent sur des listes de mots courtes, amenant ainsi à un effet de mémorisation important, majoré d'autant plus chez les personnes utilisant ces tests fréquemment (orthophonistes). De plus, chacun de ces tests utilisent des mots (et non des pseudo-mots), si bien que l'auditeur est en capacité de restaurer les séquences distordues à l'aide d'une restauration lexicale (Lalain et al., 2019). Il en résulte une surévaluation de l'intelligibilité et une mauvaise évaluation du handicap communicationnel, qui va de ce fait impliquer une rééducation orthophonique inadaptée au déficit. De plus, tous ces tests reposant sur l'évaluation perceptive par un ou plusieurs auditeurs, ils peuvent manquer d'objectivité et de reproductibilité : un patient évalué à un moment t par un auditeur sera possiblement évalué d'une autre manière par un autre auditeur avec le même test. Notre problématique est ainsi liée au manque d'outil validé pour une évaluation objective, fine et reproductible de l'intelligibilité., gênant la prise en charge orthophonique et le suivi du patient traité pour un cancer ORL.

Nos connaissances quant à la manière précise dont les traitements altèrent l'intelligibilité restent également peu satisfaisantes. Certes, la parole est altérée (Borggreven et al., 2005), mais comment ? En fonction de quelles caractéristiques de la tumeur ? Avec quel traitement ? En fonction des études, l'intelligibilité est plus altérée lorsque les patients ont bénéficié de chirurgie, et encore plus avec la mise en place d'un lambeau pour combler la perte de substance (Balaguer et al., 2019) ou associée à de la radiothérapie (Borggreven et al., 2005). La taille de la tumeur (Balaguer et al., 2019 ; Borggreven et al., 2005), sa localisation à la cavité buccale (Borggreven et al., 2007) ou à la langue mobile (Borggreven et al., 2005) et le volume de résection tumorale (Kreeft et al., 2009) semblent être également corrélés à l'altération de l'intelligibilité. Certains auteurs se sont portés sur les segments de parole distordus, par exemples les consonnes. Borggreven et al. (2005) retrouvaient des taux d'erreurs consonantiques faibles pour des cancers de la base de langue, et très élevés pour l'amygdale et la langue mobile. Pauloski et al. (1998) obtenaient un pourcentage de phonèmes consonantiques

compris inversement corrélé au volume de résection tumorale, en détaillant quels phonèmes étaient les plus altérés en fonction de la localisation de la tumeur. Enfin, la nasalité a également été étudiée puisque Borggreven constatait une altération de celle-ci chez 67% des patients un an après traitement, en particulier pour les tumeurs du voile du palais ou de l'amygdale. Malgré ces quelques études, il persiste encore des zones d'ombre, comme pour les voyelles.

Devant ce manque de connaissance quant aux effets d'une tumeur et de ses traitements sur l'intelligibilité, sans outil validé pour une évaluation fiable de celle-ci, deux projets financés ont vu le jour.

I.4. Projets C2SI et RUGBI

Ce manque d'outil fiable et objectif pour l'évaluation de l'intelligibilité rend donc difficile l'appréciation fine de l'altération de celle-ci que peut entraîner un cancer des VADS. Ceci a mené des cliniciens et des chercheurs à élaborer des projets de recherche afin de répondre à cette question. Le projet C2SI, pour Carcinologic Speech Severity Index, a été mené de 2014 à 2018 et a obtenu un financement par l'Institut National du Cancer (INCa). Dirigé par le Professeur Virginie Woisard, il avait pour but le développement d'un indice de sévérité des troubles de production de la parole pour la mesure de l'impact des traitements des cancers de la cavité buccale et de l'oropharynx sur la production de la parole, à la fois par des méthodes perceptives et par traitement automatique de la parole (Astesano et al., 2018). Un corpus de parole a été constitué pour ce projet et incluait des enregistrements de 85 patients traités pour un cancer de la cavité buccale ou de l'oropharynx et de 41 sujets contrôles, ayant effectué plusieurs tâches (parole spontanée, description de photos, lecture…). Pour ce projet, le Laboratoire Parole et Langage a notamment développé un nouvel outil d'évaluation de l'intelligibilité : la tâche de décodage acoustico-phonétique, ou encore DAP (Ghio et al., 2018). Cette tâche repose sur la production de pseudo-mots, c'est-à-dire des mots dénués de sens mais respectant les règles phonotactiques du français (ex : stoumi, clorton, fiman…). Le locuteur produit 52 pseudo-mots, choisis aléatoirement parmi une liste de 89346 pseudo-mots. Des auditeurs sont ensuite chargés de transcrire orthographiquement le mot qu'ils avaient perçu, et deux types de calcul de distance sont réalisés : la distance globale qu'il existe entre le pseudo-mot attendu et le mot transcrit, et la distance locale entre chaque phonème de chaque pseudo-mot et le phonème transcrit correspondant (Fig.5). La validité de cette tâche est déjà admise. On peut dire effectivement que l'utilisation de pseudo-mots et en grand nombre permet de réduire la prédictibilité des items et donc la capacité des auditeurs à restaurer les séquences de parole altérées (réduction des effets top-down de mémorisation et de restauration lexicale dont nous avons précédemment parlé). Lalain et al. (2019) montrent en effet une abolition de l'effet

de familiarité des auditeurs à la tâche, en comparaison au test BECD classique. Cette tâche est également corrélée au jugement subjectif de la sévérité obtenue auprès d'experts (Ghio et al., 2020), et est discriminante entre patients et contrôles puisqu'il existe une différence significative de scores entre les deux groupes (Ghio et al., 2018 ; Lalain et al., 2020).

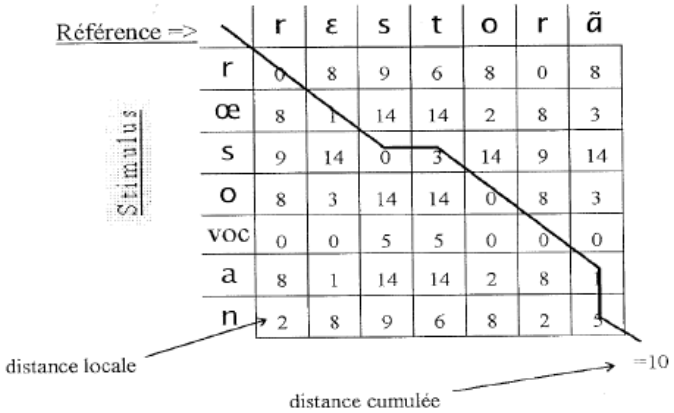

Figure 5 : Exemple de calcul de distance entre un mot cible [rɛstorã] et un mot retranscrit [roeso voc an]. La distance locale entre les phonèmes /r/ et /r/ est de 0 trait, ou encore de 1 trait d'écart entre /ɛ/ et /oe/. La distance cumulée entre la cible et le transcrit est de 10 traits d'écart.

Le projet RUGBI fait suite au projet C2SI. Débuté en 2018, il est actuellement en cours et a obtenu un financement ANR. Il a pour objectif la recherche d'unités linguistiques pertinentes pour améliorer la mesure de l'intelligibilité de la parole altérée par des troubles de production pathologique, pour aboutir à une modélisation automatique du traitement de la parole, utilisable par le clinicien. En d'autres termes, suite au projet C2SI, le projet RUGBI tente d'identifier chez ces patients les segments altérés, ou plus généralement les macro-classes concernées, afin de dégager ceux ou celles qu'il serait pertinent d'évaluer, pour créer une méthode d'évaluation encore plus ciblée et efficace de l'intelligibilité. Notre travail se base donc sur le corpus de parole du projet C2SI et s'intègre parfaitement dans le cadre du projet RUGBI.

II. PROBLEMATIQUES

La problématique de ce mémoire est donc double. Elle porte d'une part sur le lien entre la taille et la région d'une tumeur et l'intelligibilité, et d'autre part sur la capacité de la tâche de DAP à établir un lien précis entre l'intelligibilité et des données anatomiques et cliniques.

Nous faisons face en effet à une imprécision des connaissances quant à la manière avec laquelle une tumeur de la cavité buccale ou de l'oropharynx altère l'intelligibilité après traitement, de manière précise. Bien qu'une tumeur de la cavité buccale semble être plus délétère sur l'intelligibilité qu'une tumeur de l'oropharynx, ceci n'est pas encore retrouvé dans toutes les études, et il existe encore un manque d'informations quant aux facteurs pouvant influencer l'intelligibilité, mais aussi sur les segments de parole qui sont les plus altérés. Nous sommes également confrontés à un manque d'outil fiable et objectif pour l'évaluation de l'intelligibilité, validé de manière consensuelle dans la communauté scientifique et utilisable de manière routinière par les cliniciens. Avec la création de la tâche de DAP, le projet C2SI tente de pallier ce problème. Ce travail de mémoire va utiliser cette tâche dans un cadre clinique, et participera à l'évaluation de sa pertinence, s'intégrant alors dans la démarche globale du projet RUGBI.

L'objectif principal de ce travail sera de mettre en lien les données du test de jugement perceptif du DAP et la taille et la localisation des tumeurs des patients, afin de déterminer comment ces deux caractéristiques jouent sur l'intelligibilité, et ce de manière globale (grâce aux scores globaux) et segmentale (grâce aux scores locaux). Y'a-t-il des segments de parole plus altérés que d'autres ? Y'en a-t-il au contraire qui résistent quelle que soit la taille ou la localisation ? C'est en répondant à cette question principale que l'on pourra évaluer la pertinence de la tâche de DAP comme une tâche adaptée à la clinique, à savoir apte à établir un lien précis entre des segments de parole altérés et des éléments anatomiques touchés par un cancer et son traitement. La mise en relation des scores d'intelligibilité avec les données cliniques devrait permettre de prédire le degré d'altération communicationnelle en fonction de la région et de la taille d'une tumeur. L'analyse plus ciblée des segments de parole devrait quant à elle permettre de confirmer ou d'infirmer nos intuitions et d'affiner le protocole de prise en charge orthophonique.

III. MATERIELS ET METHODES

Dans cette étude, nous nous appuyons sur les données du corpus C2SI (Woisard et al., 2020), et plus précisément sur les données d'un test de jugement perceptif de l'intelligibilité (la tâche de DAP) et les métadonnées des patients. Avant de décrire le traitement et l'analyse de ces données, il est important d'expliquer comment celles-ci ont été obtenues.

III.1. Corpus

Le corpus C2SI (Woisard et al., 2020) est constitué des enregistrements de 126 locuteurs au total : 41 sujets sains et 85 sujets traités pour un cancer de la cavité buccale ou de l'oropharynx. Ils ont été enregistrés sur plusieurs tâches de production, comme de la lecture, de la description d'image, de la parole spontanée, la tâche de DAP mais aussi des tâches de prosodie. Pour être inclus, les patients devaient avoir été traités à l'Institut Universitaire du Cancer à Toulouse pour un cancer de la cavité buccale ou de l'oropharynx, toutes tailles confondues (T1 à T4). Le traitement pouvait être une chirurgie et/ou une radiothérapie et/ou une chimiothérapie. Le délai post-traitement devait être de 6 mois minimum, et le patient devait être en rémission clinique afin que la parole soit la plus stable possible. Les critères d'exclusion étaient une autre atteinte responsable de trouble de la parole (atteinte neurologique type accident vasculaire cérébral, ou encore bégaiement), et des troubles visuels, auditifs ou cognitifs incompatibles avec le bon déroulement de la tâche. Étaient recueillies les données suivantes : âge, sexe, région atteinte par le cancer (cavité buccale ou oropharynx), localisation précise (langue, plancher buccal, trigone rétromolaire, mandibule pour la cavité buccale ; amygdale, base de langue, voile du palais, « oropharynx » si plusieurs de ces localisations étaient atteintes pour l'oropharynx), stades T et N de la tumeur, délai post-traitement, traitement reçu, type de chirurgie, type de lambeau s'il y avait eu reconstruction, dose de radiothérapie, récidive ou pas. Tous les patients ont été inclus et enregistrés dans le service d'oncoréhabilitation de l'Oncopole à Toulouse. Les sujets contrôles quant à eux ont été recrutés en respectant les critères d'exclusion valables pour les patients, et devaient avoir un bon niveau en lecture. Ils ont été enregistrés à Toulouse, dans le Vaucluse et au LPL.

Pour notre travail, nous avons utilisé les métadonnées cliniques des patients (comme l'âge, la région atteinte par la tumeur, la localisation précise, la taille de la tumeur, etc…) et les données du test perceptif du DAP de ces mêmes patients.

Au total, la population du corpus était constituée de 85 patients et 41 sujets contrôles. 65 étaient des hommes (51.6%) et 61 des femmes (48.4%) : 46 hommes et 39 femmes parmi

les patients, et 19 hommes et 22 femmes parmi les sujets contrôles. L'âge moyen était de 64.2 ans (écart type 10.7 ans). Parmi les patients, 33 avaient été traités pour un cancer de la cavité buccale (38.8%), 52 pour un cancer de l'oropharynx (61.2%) (Fig.6). Les localisations plus précises étaient réparties comme telle : 16 planchers buccaux, 10 langues, 6 rétromolaires, 5 mandibules pour la cavité buccale, et 27 amygdales, 15 bases de langue, 11 oropharynx et 4 voiles du palais pour l'oropharynx. On comptait 10 tumeurs T1, 36 tumeurs T2 (soit 46 tumeurs T1T2), 12 tumeurs T3 et 27 tumeurs T4 (soit 39 tumeurs T3T4) (Fig.7). Le groupe cavité buccale regroupait 17 tumeurs T1T2 (51%) et 16 tumeurs T3T4 (48%), le groupe oropharynx regroupait 29 tumeurs T1T2 (55%) et 23 tumeurs T3T4 (44%). 89% des patients avaient bénéficié d'une chirurgie, 53% d'une chimiothérapie, 94% d'une radiothérapie (de manière associée ou pas). La durée post-traitement était en moyenne de 63.4 mois (écart type 54.2 mois).

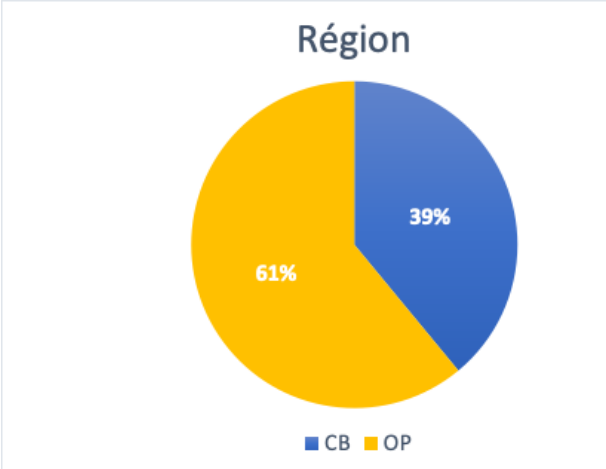

Figure 6 : Répartition des patients du corpus en fonction de la région atteinte par la tumeur.

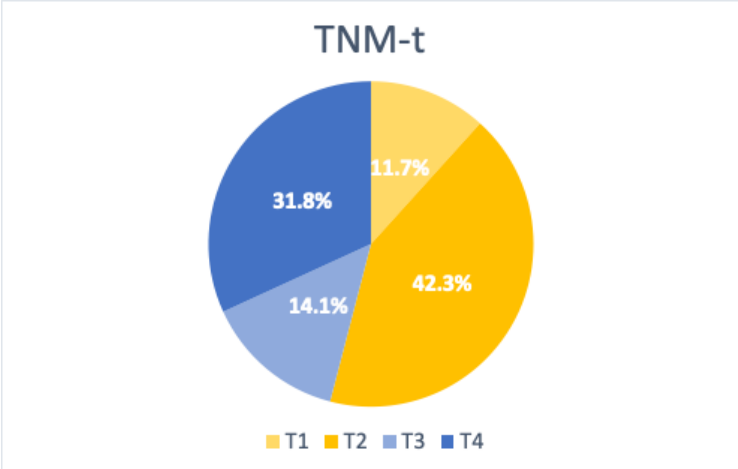

Figure 7 : Répartition des patients du corpus en fonction du stade TNM-t (taille) de leur tumeur. Le regroupement des tailles T1-T2 et des tailles T3-T4 est visuellement évident.

III.2. Un test de jugement perceptif de l'intelligibilité : la tâche de DAP

Le but en créant la tâche de DAP était de supprimer les limites rencontrées dans les tests classiques d'évaluation de l'intelligibilité, qui mènent à une surévaluation de celle-ci (Lalain et al., 2020). Chaque auditeur est en effet capable de restaurer des séquences distordues lorsqu'il est confronté à un énoncé altéré, grâce aux processus top-down, c'est-à-dire aux informations non dépendantes du signal de parole mais plutôt au contexte et aux connaissances de l'auditeur. Lorsque ce dernier perçoit un discours altéré, ces processus vont agir comme des mécanismes de compensation afin d'optimiser au maximum l'intelligibilité du message. On retrouve notamment l'effet de la lexicalité : l'exemple bien connu est celui de la distinction parfois

difficile entre /t/ et /d/. Devant la séquence [aʃ], l'auditeur percevra plutôt un /t/ (puisqu'il possède dans son lexique le mot « tâche »), alors qu'il le percevra /d/ devant la séquence [isk] (en référence au mot « disque »). La fréquence lexicale intervient aussi puisque les mots fréquents sont plus facilement identifiés. L'auditeur est donc capable d'effectuer une restauration phonémique lorsqu'un énoncé est dégradé. La majeure partie des tests utilisés en routine pour évaluer l'intelligibilité utilisant des mots connus de tous, ces mécanismes de compensation sont donc rencontrés et altèrent l'évaluation. L'autre problème rencontré est celui de l'effet de mémorisation : les listes de mots utilisés pour les tests étant courtes, les cliniciens habitués à les utiliser quotidiennement viennent à les connaître presque par cœur, si bien qu'ils peuvent restaurer les séquences altérées par mémorisation des items.

Pour lutter contre cela, il a été proposé l'utilisation de pseudo-mots, c'est-à-dire des items respectant les règles phonotactiques du français mais dépourvus de sens (Lalain et al., 2020), et ce en grand nombre. Ceux-ci ont été générés au cours du projet C2SI à l'aide d'un module de la base de données « Lexique.org ». Chaque pseudo-mot respectait la structure $C(C)_1V_1C(C)_2V_2$, où $C(C)_i$ est une consonne isolée ou un groupe consonantique. La fréquence d'apparition des phonèmes dans la langue en fonction de leur position ainsi que la fréquence d'apparition des groupes consonantiques ont été prises en compte. Une multitude de pseudo-mots a été obtenue, par exemple stoumo, vurtan, stolain, blanga, staji … Après avoir éliminé les items qui correspondaient à des vrais mots ou des débuts de mots réels, les auteurs ont obtenu un total de 89346 pseudo-mots. Suite à un tirage au sort plusieurs listes de 52 pseudo-mots ont été générées, toutes différentes et phonétiquement équilibrées (Fig.8). C'est par ce grand nombre de pseudo-mots et de combinaisons possibles que la neutralisation de l'effet d'apprentissage par l'auditeur est faisable. L'effet de restauration lexicale est quant à lui aboli par l'utilisation de pseudo-mots (Rebourg et al., 2020)

crampant	fevo	quinfant	flaspou	plouniant
troucha	suptu	vabla	baillu	ratri
rougli	nougu	touflant	griti	ninrtin
dibro	yango	zucrou	quebo	gavi
pufru	scuchu	psoussa	chanjin	bijo
blouillu	fampsi	madin	lupou	tanli
niascu	pimprant	climbou	storquin	brori
chansin	jindou	spucou	glima	prinrmo
gomou	droto	yezant	vefin	zelin
Jezant	dorba	nioniou	mera	frina
Lina	siqui			

Figure 8 : Exemple de liste de 52 pseudo-mots.

Ce test a été utilisé dans les projets C2SI et RUGBI pour évaluer l'intelligibilité. Nous allons voir qu'il est constitué de trois étapes différentes.

III.3. Évaluer l'intelligibilité avec la tâche de DAP

III.3.1. Phase de production

La première étape de la tâche de DAP consiste à enregistrer les locuteurs sur la production des listes de 52 pseudo-mots. Au cours d'une de leurs consultations de suivi à l'Institut Universitaire du Cancer à Toulouse, les patients sont installés dans une cabine anéchoïque munie d'un écran, d'un microphone Neumann TLM 102 Cardioid Condenser relié à un enregistreur digital Fostex et d'un filtre anti-pop. La fréquence d'enregistrement est fixée à 48 kHz. Puis 52 pseudo-mots leur sont présentés successivement. Chaque stimulus (chaque pseudo-mot) est présenté selon une double modalité : visuelle et auditive. Le locuteur est placé devant l'écran sur lequel s'affiche une version orthographiée du pseudo-mot à prononcer et entend de manière synchrone une version orale. Celle-ci était initialement une voix humaine préalablement enregistrée, puis devant des difficultés de compréhension et le grand nombre de pseudo-mots les auteurs ont opté pour une voix de synthèse (créée à l'aide de la synthèse vocale Voxygen www.voxygen.fr). Cette double modalité visuelle et auditive permet de limiter les erreurs de lecture ainsi que de contrer les difficultés auditives et attentionnelles éventuelles. Chaque locuteur doit prononcer ces 52 pseudo-mots et est enregistré.

III.3.2. Phase de perception

Une fois les enregistrements effectués, il s'agit ensuite de les faire écouter aux auditeurs afin qu'ils effectuent une transcription orthographique des items perçus. Ces derniers sont français natifs, ne doivent pas présenter de troubles auditifs ou visuels et avoir une bonne maîtrise de l'orthographe.

Une étape préalable consiste à segmenter les enregistrements des patients afin de créer un fichier audio par pseudo-mot. Pour cela, le logiciel Praat et son extension EasyAlign sont utilisés. Les stimuli sonores obtenus sont ensuite tous randomisés, si bien que chaque auditeur va écouter des pseudo-mots produits par plusieurs locuteurs différents. Et chaque pseudo-mot produit par un locuteur est écouté par 3 auditeurs différents.

La tâche de perception, qui s'est déroulée au Centre d'Expérimentation de la Parole (http://cep.lpl-aix.fr/) au Laboratoire Parole et Langage à Aix-en-Provence, est réalisée grâce à la plateforme Perceval et son application Lancelot. Les auditeurs sont munis d'un casque audio, d'un écran et d'un clavier (Fig.9). La consigne suivante leur est donnée : « Vous allez entendre des pseudo-mots. Un pseudo-mot est une combinaison de sons de la langue française qui n'a

pas de signification (ex : gloutu). En respectant les règles de l'orthographe du français, vous devrez transcrire ce que vous entendrez. Certaines prononciations seront difficiles à identifier mais dans tous les cas, vous devrez proposer une transcription. ». Un stimulus sonore leur est diffusé, et ils doivent écrire sur le clavier ce qu'ils entendent. Ils peuvent écouter chaque stimulus au maximum 3 fois.

Figure 9 : Exemple de déroulement de phase perceptive. Chaque auditeur est muni d'un casque audio et placé devant un écran.

Enfin, à l'issue du test un fichier est généré par Perceval, qui comprend les cibles attendues (qui devaient être prononcées par les locuteurs) et leurs transcriptions phonétiques au format Sampa, les réponses orthographiées par les auditeurs et leurs versions phonétiques au format Sampa. Ce sont ces deux transcriptions qui sont ensuite comparées pour le calcul du score.

III.3.3. Calcul du score d'intelligibilité

Une fois les transcriptions par les auditeurs effectuées, un certain nombre d'opérations de traitement des données ont été effectuées (pour plus de détail, voir Lalain et al., 2020) afin de calculer, pour chaque locuteur, un score d'intelligibilité DPP (PPD) ou score de Déviation Phonologique Perçue. Ce score est calculé sur la base d'une matrice de coût créée par l'équipe du LPL au cours du projet C2SI permettant de mesurer la distance qu'il existe entre chaque pseudo-mot attendu et le pseudo-mot retranscrit par l'auditeur. Ghio et collègues (Ghio et al., 2018 ; Lalain et al., 2020) ont en effet établi une matrice de coût entre phonèmes, c'est-à-dire un tableau représentant le degré de dissimilitude entre eux (Fig.10). Cette matrice repose sur la théorie des traits selon Jakobson (Jakobson et al., 1951), qui établit que chaque phonème peut être décomposé en un certain nombre de traits qui permet de les distinguer. En suivant et en complétant la décomposition en traits des voyelles et consonnes par Chomsky et Halle (1968), Ghio a établi de manière binaire qu'un trait était présent ou absent (1 ou 0 respectivement) chez un phonème, et a obtenu ainsi les matrices suivantes pour les voyelles et les consonnes, où apparait le nombre de traits d'écart entre chaque phonème.

	a	i	u	o	e	y	ø	ɛ	ɔ	œ	ɑ̃	ɛ̃	ɔ̃	œ̃	Ê	Ô	Û	µ	&
a	0	3	3	2	2	4	3	1	1	2	1	2	2	3	1	1	2	2	1
i	3	0	2	3	1	1	2	2	4	3	4	3	5	4	1	3	2	3	1
u	3	2	0	1	3	1	2	4	2	3	4	5	3	4	3	1	2	4	2
o	2	3	1	0	2	2	1	3	1	2	3	4	2	3	2	0	1	3	1
e	2	1	3	2	0	2	1	1	3	2	3	2	4	3	0	2	1	2	0
y	4	1	1	2	2	0	1	3	3	2	5	4	4	3	2	2	1	3	1
ø	3	2	2	1	1	1	0	2	2	1	4	3	3	2	1	1	0	2	0
ɛ	1	2	4	3	1	3	2	0	2	1	2	1	3	2	0	2	1	1	0
ɔ	1	4	2	1	3	3	2	2	0	1	2	3	1	2	2	0	1	2	1
œ	2	3	3	2	2	2	1	1	1	0	3	2	2	1	1	1	0	1	0
ɑ̃	1	4	4	3	3	5	4	2	2	3	0	1	1	2	2	2	3	1	2
ɛ̃	2	3	5	4	2	4	3	1	3	2	1	0	2	1	1	3	2	0	1
ɔ̃	2	5	3	2	4	4	3	3	1	2	1	2	0	1	3	1	2	1	2
œ̃	3	4	4	3	3	3	2	2	2	1	2	1	1	0	2	2	1	0	1
Ê	1	1	3	2	0	2	1	0	2	1	2	1	3	2	0	2	1	1	0
Ô	1	3	1	0	2	2	1	2	0	1	2	3	1	2	2	0	1	2	1
Û	2	2	2	1	1	1	0	1	1	0	3	2	2	1	1	1	0	1	0
µ	2	3	4	3	2	3	2	1	2	1	1	0	1	0	1	2	1	0	1
&	1	1	2	1	0	1	0	0	1	0	2	1	2	1	0	1	0	1	0

	p	t	k	b	d	g	f	s	S	v	z	Z	m	n	N	l	R
p	0	1	2	1	2	3	1	2	3	2	3	4	3	4	5	3	4
t	1	0	1	2	1	2	2	1	2	3	2	3	4	3	4	2	3
k	2	1	0	3	2	1	3	2	1	4	3	2	5	4	3	3	4
b	1	2	3	0	1	2	2	3	4	1	2	3	2	3	4	2	3
d	2	1	2	1	0	1	3	2	3	2	1	2	3	2	3	1	2
g	3	2	1	2	1	0	4	3	2	3	2	1	4	3	2	2	3
f	1	2	3	2	3	4	0	1	2	1	2	3	4	5	6	4	3
s	2	1	2	3	2	3	1	0	1	2	1	2	5	4	5	3	2
S	3	2	1	4	3	2	2	1	0	3	2	1	6	5	4	4	3
v	2	3	4	1	2	3	1	2	3	0	1	2	3	4	5	3	2
z	3	2	3	2	1	2	2	1	2	1	0	1	4	3	4	2	1
Z	4	3	2	3	2	1	3	2	1	2	1	0	5	4	3	3	2
m	3	4	5	2	3	4	4	5	6	3	4	5	0	1	2	2	3
n	4	3	4	3	2	3	5	4	5	4	3	4	1	0	1	1	2
N	5	4	3	4	3	2	6	5	4	5	4	3	2	1	0	2	3
l	3	2	3	2	1	2	4	3	4	3	2	3	2	1	2	0	1
R	4	3	4	3	2	3	3	2	3	2	1	2	3	2	3	1	0

Figure 10 : Matrices de coût des voyelles (à gauche) et des consonnes (à droite). Ê, Û et Ô étant des archi-phonèmes correspondant respectivement à {e, ɛ}, {oe, ø} et {o, ɔ}. (Ghio et al., 2018)

Ces matrices permettent d'intégrer les phénomènes d'insertion, d'élision et de substitution d'unités. Ce calcul de distance de Levenshtein porte ici sur les phonèmes, et non pas sur les formes orthographiques des mots. Ceci permet d'obtenir des nuances plus subtiles que la simple comparaison binaire « similaire/différent » que l'on a classiquement avec la comparaison des formes orthographiques. La comparaison des phonèmes permet d'être plus spécifique. Par exemple une confusion entre une consonne et une voyelle sera plus importante et aura plus de poids qu'une confusion entre deux voyelles.

Grâce à l'utilisation de ces matrices de coût et à un algorithme de Wagner Fisher (Fig.11), nous obtenons une distance locale pour chaque couple phonème attendu – phonème transcrit, et une distance cumulée pour le couple pseudo-mot attendu – pseudo-mot transcrit. Ce que nous appelons score PPD (pour *perceived phonologic deviation*, ou déviation phonologique perçue) est en fait la distance cumulée entre tous ces phonèmes divisée par le nombre de phonème constituant le pseudo-mot, il s'exprime en nombre de traits altérés par phonème. On considère que l'intelligibilité est altérée à partir du score PPD seuil de 0.6 trait altéré par phonème (Ghio et al., 2020).

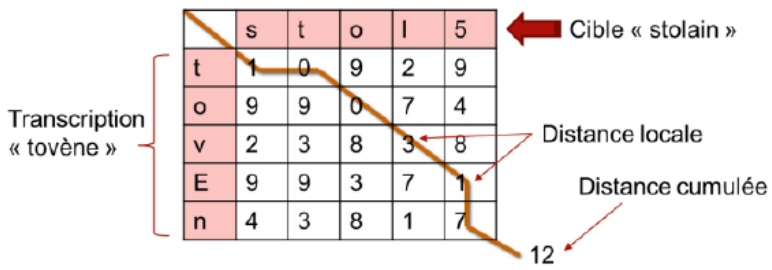

Figure 11 : Calcul de distance entre deux pseudo-mots à l'aide d'un algorithme de Wagner-Fisher. La cible attendue était « stolain », l'auditeur a retranscrit « tovène ».

Un exemple de calcul de distance entre un pseudo-mot attendu et un pseudo-mot retranscrit est visible sur la figure 11 : le locuteur devait produire le pseudo-mot *stolain*, l'auditeur a transcrit orthographiquement *tovène*. Après transformation phonétique, on calcule la distance entre chaque phonème, elle est par exemple de 0 trait entre /o/ et /o/, et de 3 traits entre /l/ et /v/. La distance cumulée entre ces deux pseudo-mots est donc de 12, soit 12 traits d'écart. Le score PPD sera ainsi égal à 12/5 = 2.4 traits altérés par phonème, puisque le pseudo-mot est constitué de 5 phonèmes. Le score global d'un patient (ou score PPD global) est obtenu en calculant la moyenne des scores de chacune de ses 52 productions.

Il faut cependant préciser que l'algorithme initial ne nous permettait pas de connaître les distances locales existantes entre chaque phonème, cela constituait une sorte de « boîte noire ». Alain Ghio, ingénieur de recherche au LPL, a donc retravaillé l'algorithme afin d'en dégager les distances locales (ou score PPD local) c'est-à-dire les distances entre chaque phonème attendu et chaque phonème transcrit correspondant pour chacun des pseudo-mots. Ce score PPD local incluait donc les phénomènes de substitution, élision ou insertion de phonème.

Nous avons donc travaillé sur deux types de scores : les scores PPD globaux, correspondant à un score général pour chaque patient et étant la moyenne des scores de ses 52 pseudo-mots produits, et les scores PPD locaux reflétant la distance existante entre chaque phonème produit et le phonème transcrit correspondant, permettant une analyse plus précise de l'altération par l'identification des phonèmes les plus (ou les moins) altérés.

III.4. Traitement des données

Une fois les données recueillies, nous avons obtenu deux types de résultats : les premiers sont les scores PPD globaux, correspondant pour chaque patient à la moyenne des scores PPD de ses 52 productions, et les deuxièmes sont les scores PPD locaux correspondant à la distance locale entre chaque phonème attendu et le phonème transcrit correspondant. La figure 12 est tirée de notre premier jeu de données regroupant toutes les métadonnées de chaque patients et leur score PPD global respectif, et la figure 13 est tirée de notre deuxième jeu de données, montrant en détail chaque production de chaque locuteur, et le score PPD local de chaque phonème de chaque pseudo-mot produit.

	A	B	C	D	E	F
1	ID	age	region	loc	tnmt	dapp1
2	QUD71_	74	2	voile	1	4.071538
3	TTP043	67	2	oropharynx	2	2.839391
4	COM100	58	2	oropharynx	3	2.517019
5	TTP044	60	2	bdl	2	2.372756
6	BEH66_	76	2	amygdale	2	2.309052
7	BOG109	65	1	langue	4	2.299306
8	BOM123	82	1	retromolaire	1	2.287372
9	NOF004	59	2	amygdale	4	2.170705
10	BOM94_	58	1	plancher	4	2.133301
11	BOP97_	81	1	retromolaire	2	2.092347

Figure 12 : Métadonnées cliniques et score PPD global. Exemple du locuteur NOF004, atteint d'un cancer de l'amygdale classé T4 (région 2 = oropharynx). Il présente un score PPD global de 2.17, c'est-à-dire une moyenne de 2.17 traits altérés par phonème sur l'ensemble de ses 52 productions

	A	B	C	D	E	F	G	H
1	Lot	group	speaker	region	listener	Expected	perceived	PPD
291	1	PATIENT	NOF004	2	AA1	l	l	0
292	1	PATIENT	NOF004	2	AA1	a	a	0
293	1	PATIENT	NOF004	2	AA1	n	n	0
294	1	PATIENT	NOF004	2	AA1	@	o	3
295	1	PATIENT	NOF004	2	AA1			

Figure 13 : Score PPD local. Toujours sur l'exemple du locuteur NOF004 : sur la production du pseudo-mot « lanan », l'auditeur a perçu « lano ». Les scores PPD locaux des phonèmes /l/, /a/ et /n/ sont égaux à 0 donc parfaits, le score local pour le phonème /ã/ perçu /o/ est de 3 traits d'écart. @ = phonème /ã/, µ = phonème /ɛ̃/

Avant de lancer toute analyse statistique, nous avons d'abord effectué un dépouillement et prétraitement des données sur le logiciel Excel, ainsi que des analyses graphiques, afin de dégager les résultats qui semblaient pertinents. Rappelons que les scores PPD reflètent une altération de l'intelligibilité : plus le score est élevé, plus l'altération est grande, avec une valeur seuil de 0.6 trait altéré par phonème.

Dans ce travail, nous avons cherché à identifier d'une part l'impact qu'ont la taille de la tumeur et sa localisation sur l'intelligibilité de manière globale, et d'autre part si certains segments sont plus altérés que d'autres. Pour les effets de la taille et de la localisation de la tumeur, nous avons d'abord travaillé sur les données PPD global. Nous avons mis en relation les scores PPD globaux des patients avec la région atteinte par le cancer (cavité buccale ou oropharynx). Les scores globaux semblaient plus élevés pour les cancers de la cavité buccale. Quand on prenait en revanche toutes les localisations une par une (mandibule, amygdale, …), il était plus difficile de conclure. Lorsque l'on comparait les scores globaux des patients en fonction du stade T (taille, de T1 à T4) de leur tumeur, aucune tendance marquante ne semblait ressortir. Nous avons donc regroupé les petites tumeurs (T1 et T2) et les grosses tumeurs (T3 et T4), et lorsque nous avons comparé ces deux nouveaux supra-groupes la différence était non équivoque : les grosses tumeurs T3T4 présentaient des scores globaux beaucoup plus élevés que les petites tumeurs T1T2 (Fig.14). Ce regroupement par taille paraît évident quand nous regardons la répartition des différentes tailles de tumeur dans le corpus (Fig.15).

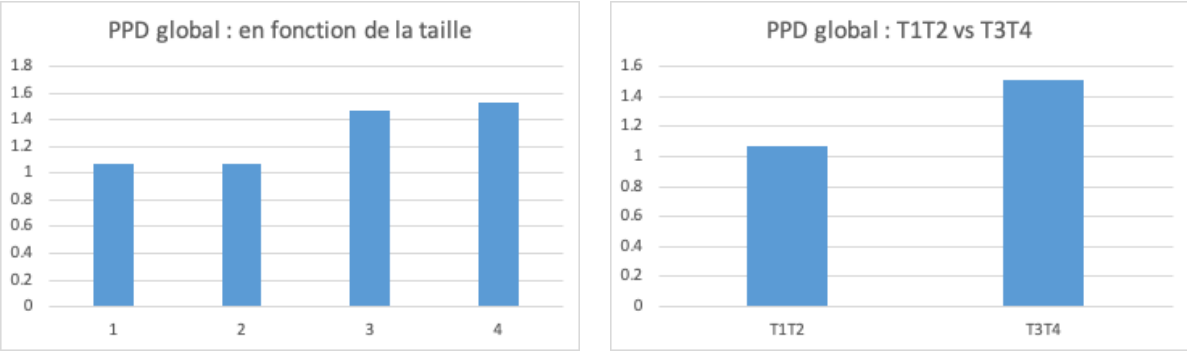

Figure 14 : PPD global en fonction de la taille de la tumeur. A gauche, PPD global en fonction de chaque taille de tumeur T1 à T4. A droite, PPD global en fonction du supra-groupe de taille de tumeur T1T2 ou T3T4.

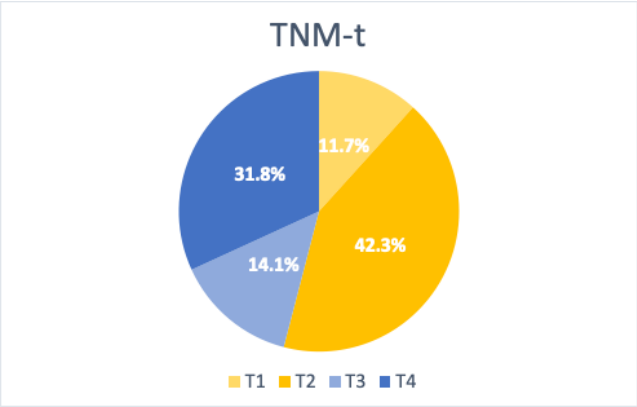

Figure 15 : Répartition des patients du corpus en fonction du stade TNM-t (taille) de leur tumeur.

Nous avons ensuite travaillé sur les données PPD local. Nous avons regardé si certains phonèmes étaient plus altérés que d'autres, en fonction de la taille et de la localisation de la tumeur. Nous n'avons regardé que les phonèmes simples, c'est-à-dire sous le format V, C ou CC, et seulement les phénomènes de substitution (en laissant de côté les phénomènes d'élision et d'insertion). Rapidement nous nous sommes aperçus que le regroupement des phonèmes selon divers facteurs (type de phonème, lieu d'articulation, …) permettait des analyses plus pertinentes. Nous avons notamment dégagé que les phonèmes de types clusters CC étaient plus altérés que les voyelles V, qui étaient elles-mêmes plus altérées que les consonnes C. Les deux voyelles nasales du corpus /ã/ et /ɛ̃/ (respectivement notées /@/ et /µ/) présentaient également des scores plus élevés que les autres voyelles orales. Sur la figure 16, nous pouvons bien voir que lorsque nous nous intéressons aux voyelles une par une, les voyelles nasales sortent du lot avec des scores plus élevés, alors que les scores de toutes les autres voyelles orales sont plus disparates. Leur regroupement semble donc pertinent. Nous avons constitué d'autres supra-groupes de voyelles en fonction de leur trait haut/bas ou antérieur/postérieur. Pour les consonnes, les regroupements qui paraissaient convaincants étaient les regroupements selon leur lieu d'articulation (labial, alvéolaire, palato-vélaire) et leur mode articulatoire (plosives, fricatives, consonnes sonnantes). Pour nos analyses, nous avons donc regroupé différents segments en prenant appui sur les oppositions phonologiques et la description articulatoire.

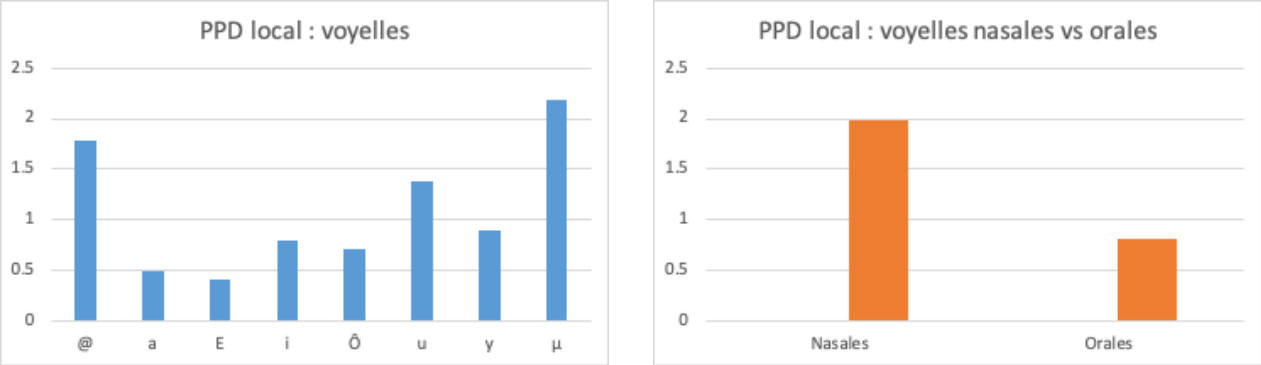

Figure 16 : Scores PPD locaux pour les voyelles. A gauche, score PPD local moyen pour chaque voyelle : les scores sont assez disparates en dehors des voyelles nasales /ã/ et /ẽ/ qui ont des scores très élevés. A droite, score PPD local moyen pour les voyelles nasales et orales.

III.5. Analyse statistique

Toutes les analyses statistiques ont été réalisées à l'aide du logiciel RStudio.

Une analyse descriptive de chaque variable a été effectuée. Un test de Shapiro a été réalisé afin de vérifier la distribution des données : pour le premier jeu de données (scores PPD globaux), la distribution des données en fonction de la taille et de la région de la tumeur n'était pas normale. Après transformation logarithmique, le test de Shapiro a été à nouveau réalisé et les distributions étaient normales. L'homogénéité des variances a été également vérifiée à l'aide des tests de Bartlett et Levene. Concernant le deuxième jeu de données (scores PPD locaux), le test de Shapiro a montré une distribution non gaussienne des données.

Ainsi, l'utilisation de tests paramétriques a donc été possible pour l'analyse des scores PPD globaux : afin d'évaluer la corrélation entre la région et la taille d'une tumeur et les scores PPD globaux, nous avons donc réalisé des ANOVA. Les variables étaient la taille de la tumeur (facteur à 2 ou 4 niveaux selon si l'on fait des regroupements) et la localisation de la tumeur (facteur à 2 niveaux). Pour l'analyse des scores PPD locaux, des tests paramétriques ont été impossibles, nous avons donc utilisé des tests de Wilcoxon, par paires de variables (consonnes vs voyelles, antérieures vs postérieures, etc…), ainsi que des tests de Cohen pour calculer la taille des effets. Le but étant de rechercher une éventuelle corrélation entre les segments altérés et la taille ou la localisation des tumeurs.

Tous les résultats étaient considérés significatifs si la p-value était inférieure à 0.05.

IV. RESULTATS

Le but de notre travail était d'analyser comment la taille d'une tumeur et sa localisation (cavité buccale ou oropharynx) influençaient l'intelligibilité, premièrement en analysant les scores PPD globaux, puis en se portant sur les scores PPD locaux afin identifier les segments de parole les plus touchés.

IV.1. Altération globale de l'intelligibilité en fonction de la taille et de la région d'une tumeur : scores PPD globaux

Une analyse de variance (ANOVA) a été réalisée afin d'identifier les effets des facteurs taille et région (variables indépendantes) sur les scores PPD globaux (variable dépendante). Le facteur taille présente 4 niveaux (T1, T2, T3, T4) ou deux niveaux si l'on effectue des regroupements (T1T2 et T3T4), et le facteur région deux niveaux (cavité buccale et oropharynx). Rappelons que la valeur seuil à partir de laquelle un score est considéré comme anormal (intelligibilité dégradée) est de 0.6 trait altéré par phonème (Ghio et al., 2020). Les résultats ne montrent pas de différence significative entre le groupe cavité buccale et le groupe oropharynx ($p = 0.3413$), on ne peut donc pas conclure sur un éventuel effet de la région sur l'intelligibilité. Concernant la taille de la tumeur, nous avons d'abord fait notre analyse avec le facteur taille à 4 niveaux. Nous ne retrouvons une différence significative d'intelligibilité qu'entre le groupe T4 et le groupe T2 ($p = 0.0334$), mais pas entre les autres groupes. Ce résultat n'est pas parfaitement satisfaisant, et connaissant les différences notables d'effectifs entre les différents groupes de taille de tumeur (Fig.15), nous avons réalisé l'analyse avec cette fois-ci les supra-groupes de tailles de tumeurs (T1T2 vs T3T4). Ce regroupement par taille est aussi retrouvé dans l'étude de Borrgreven et al. (2005), probablement pour des raisons d'effectifs également (35 tumeurs T2 vs 41 tumeurs T3 et 3 tumeurs T4) . Nous obtenons cette fois-ci une différence franchement significative entre les deux groupes : comme on peut le voir sur la figure 17, les patients ayant une grande tumeur classée T3T4 ont des scores PPD globaux plus élevés que les patients avec des petites tumeurs T1T2.

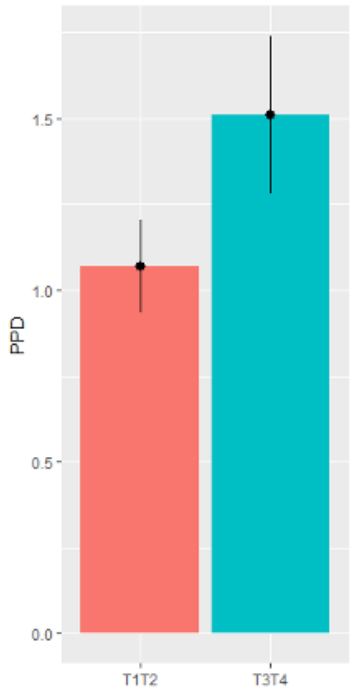

Figure 17 : Scores PPD globaux en fonction de la taille de la tumeur (petites tumeurs T1T2 vs grosses tumeurs T3T4)

Les patients avec une grande tumeur T3T4 sont donc plus altérés que les patients avec une petite tumeur, cette différence est significative ($p = 0.001$). Il y a donc un effet de la taille de la tumeur sur l'altération de l'intelligibilité.

IV.2. Altération précise des segments de parole : scores PPD locaux

Nous nous sommes ensuite portés sur les scores PPD locaux. L'idée était de dégager quels phonèmes étaient le plus particulièrement touchés, et à l'inverse ceux qui ne l'étaient pas ou peu. Nous avons d'abord reformaté nos données en enlevant les phonèmes /ɲ/ et /j/, leurs valeurs étant aberrantes et posant problème dans les analyses. Ces phonèmes seront analysés au cours de travaux futurs, notre travail de mémoire étant en effet un premier travail de dépouillement des données. Les erreurs retrouvées et les questions posées par celui-ci seront traitées par l'équipe du projet RUGBI prochainement.

Dans un premier temps, nous avons cherché à savoir quelle famille de segments était la plus touchée : les voyelles, les consonnes ou alors les clusters de consonnes. Les résultats ont montré des scores locaux plus élevés pour les clusters de consonnes, puis pour les voyelles et enfin pour les consonnes (Fig.18). Les patients sont donc beaucoup plus altérés pour la prononciation des clusters de consonnes, par rapport aux voyelles puis aux consonnes simples. La différence entre chaque famille de phonème était significative ($p < 2.2^{-16}$).

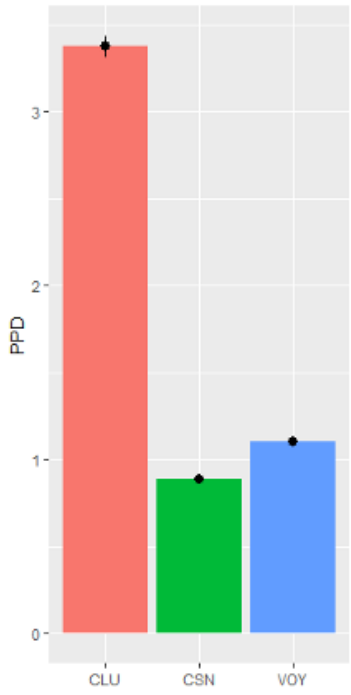

Figure 18 : Scores PPD locaux en fonction du type de phonème. CLU = clusters de consonnes, CSN = consonnes, VOY = voyelles.

Nous avons ensuite analysé les voyelles seules. Une première analyse descriptive, visible sur le graphique suivant (Fig.19), a permis de dégager rapidement la tendance suivante : les voyelles nasales présentent clairement des scores PPD locaux plus élevés que les autres voyelles. Les nasales sont suivies par les voyelles orales /u/ et /y/ qui sont des voyelles hautes. Mais le /ɔ̃/ et le /u/ sont aussi des voyelles postérieures. Ainsi nous avons décidé de regrouper les voyelles selon différents traits (nasales/orales, hautes/moyennes/basses, antérieures/postérieures) afin de les examiner en tant de macroclasses.

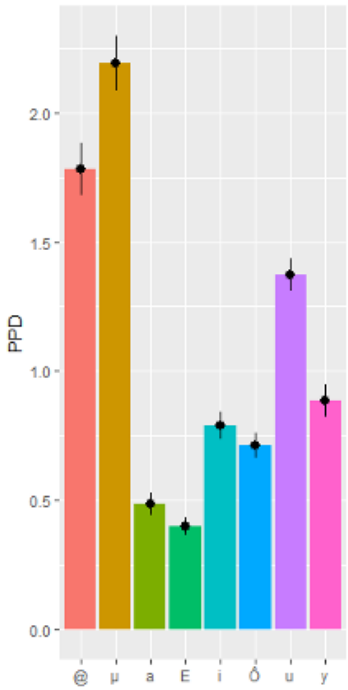

Figure 19 : Scores PPD locaux moyens pour chaque voyelle. Scores PPD locaux moyens pour chaque voyelle.
@ = /ɔ̃/, µ = /ɛ̃/, E = archiphonème /e, ɛ, oe, ø/, Ô = archiphonème /o, ɔ/

Sur la figure 20, nous voyons que le voyelles nasales /ɔ̃/ et /ɛ̃/ présentent des scores plus élevés que les voyelles orales /a, E, i, Ô, u, y/. Les patients sont donc moins intelligibles sur les voyelles nasales que sur les voyelles orales, cette différence est significative ($p < 2.2^{-16}$), et l'effet d de Cohen est modéré (d = 0.586). Nous avons testé l'effet de la région sur ces données. Comme nous pouvons le voir sur la figure 21, on s'aperçoit que les voyelles nasales ont des scores plus élevés dans le groupe oropharynx par rapport au groupe cavité buccale, et les voyelles orales ont des scores plus hauts dans le groupe cavité buccale par rapport au groupe oropharynx. Il y a donc un effet de la région sur les voyelles nasales et orales : les premières seront encore plus altérées avec une tumeur de l'oropharynx, et les deuxièmes seront encore plus altérées avec une tumeur de la cavité buccale.

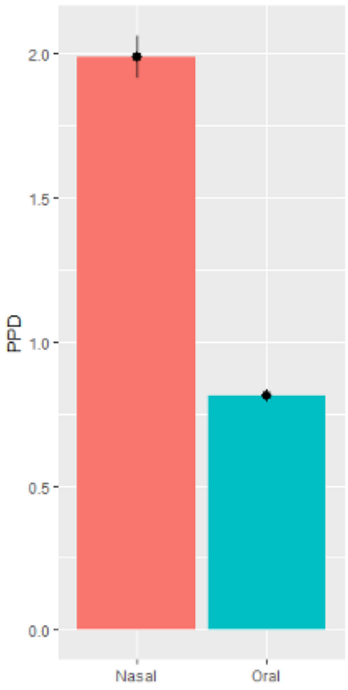

Figure 20 : Scores PPD locaux pour les voyelles nasales et orales.

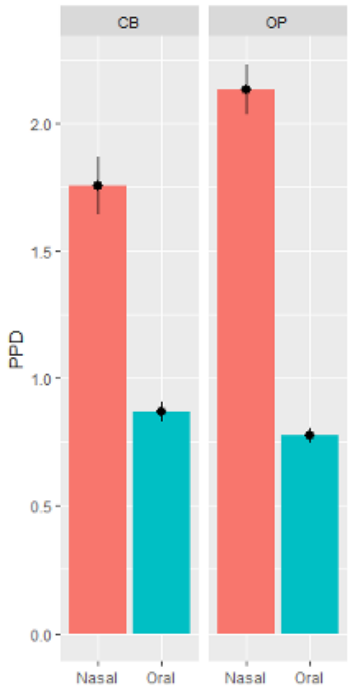

Figure 21 : Scores PPD locaux des voyelles nasales et orales en fonction de la région de la tumeur. Les voyelles nasales ont des scores plus élevés dans le groupe oropharynx, et les voyelles orales ont des scores plus élevés dans le groupe cavité buccale. Une tumeur de l'oropharynx altère donc plus l'intelligibilité des voyelles nasales qu'une tumeur de la cavité buccale.

Nous nous sommes ensuite intéressés aux voyelles orales hautes /i, u, y/, moyennes /E, Ô/ et basses /a/, en ne prenant plus en compte les voyelles nasales. Nous constatons des scores plus élevés pour les voyelles orales hautes, puis pour les voyelles orales moyennes et enfin basses (Fig.22). Encore une fois cette différence est significative ($p < 2.2^{-16}$), les voyelles orales hautes sont donc moins intelligibles chez nos patients que les voyelles orales moyennes et basses.

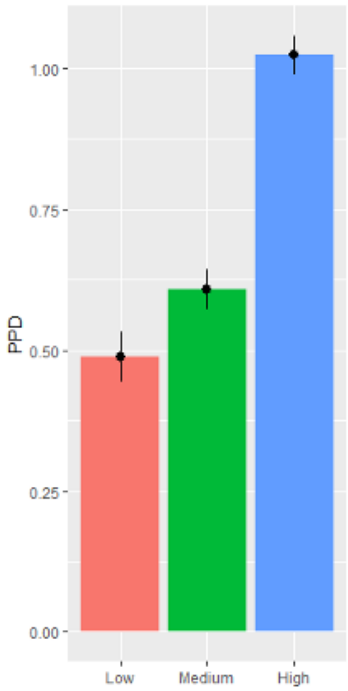

Figure 22 : Scores PPD locaux pour les voyelles basses (low), moyennes (medium) et hautes (high).

Enfin nous avons cherché une différence entre les voyelles orales selon leur réalisation antéro-postérieure. Nous avons décidé de ne pas prendre en compte /a/, l'aperture étant tellement importante pour ce phonème que la notion de la localisation antéro-postérieure devient inadaptée. La figure 23 montre des scores PPD plus élevés pour les voyelles orales postérieures /Ô, u/ par rapport aux voyelles orales antérieures /i, E, y/. Les patients sont donc plus altérés pour la prononciation des voyelles orales postérieures que pour les voyelles orales antérieures, cette différence est significative ($p < 2.2^{-16}$). En recoupant ces données avec le facteur région, nous pouvons voir que les voyelles postérieures sont encore plus altérées dans le groupe cavité buccale par rapport au groupe oropharynx (Fig.24), avec une différence significative ($p < 0.01$). La région de la tumeur a donc un impact significatif sur l'intelligibilité des voyelles postérieures, celles-ci étant altérées dans les deux groupes mais de manière plus importante en cas de tumeur de la cavité buccale.

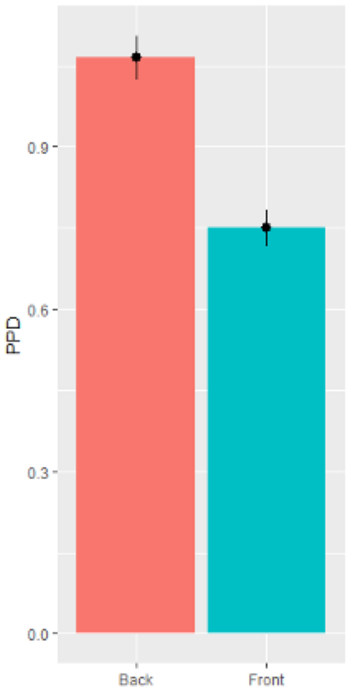

Figure 23 : Scores PPD locaux pour les voyelles postérieures (back) et antérieures (front).

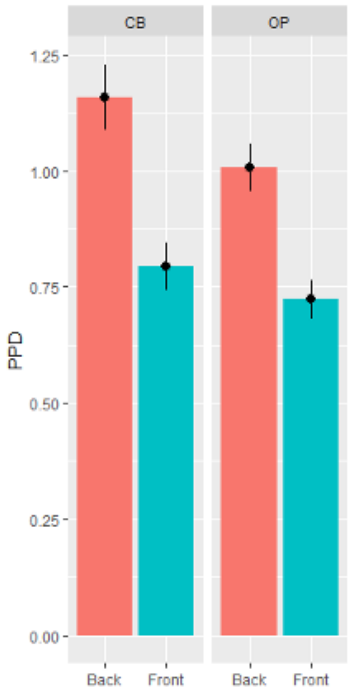

Figure 24 : Effet de la région sur les scores PPD locaux des voyelles postérieures et antérieures. Elles présentent des scores plus élevés, notamment les postérieures, dans le groupe cavité buccale.

Enfin, nous nous sommes portés sur les consonnes. Ce travail s'est encore une fois organisé autour de macro-classes de consonnes, puisque comme nous pouvons le voir sur la figure 25 les scores PPD de chaque consonne prise isolément sont très disparates, et les analyser une par une aurait manqué de puissance et de cohérence. Seules les consonnes /k/ et /g/ sortent clairement du lot avec des scores franchement élevés.

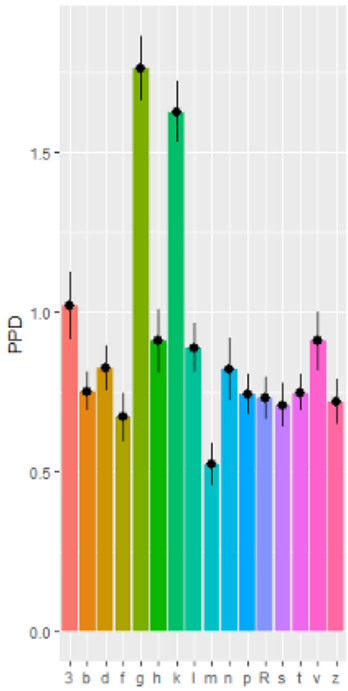

Figure 25 : Scores PPD locaux pour chaque consonne. Nous pouvons voir des scores très disparates, avec néanmoins les consonnes palatales /k/ et /g/ qui semblent présenter des scores bien plus élevés que toutes les autres consonnes. h = /ʃ/, 3 = /ʒ/

Nous avons choisi de regrouper les consonnes premièrement en fonction de leur lieu d'articulation, facteur à 3 niveaux : labial /p, b, f, v, m/, alvéolaire /t, d, s, z, n, l/ et palato-vélaire /ʒ, ʃ, k, g, R/. Les scores locaux des macro-classes constituées ont été comparés, et nous pouvons voir que les consonnes palato-vélaires ont effectivement des scores plus élevés que les consonnes labiales et alvéolaires (Fig.26). Les patients sont donc moins intelligibles sur la prononciation des consonnes palatales que sur celle des autres types de consonnes. Cette différence est significative ($p < 2.2^{-16}$). La taille de l'effet de Cohen d entre les scores des consonnes palatales et les consonnes non palatales est petit (d = -0.2). Deuxièmement, nous avons choisi de faire un regroupement en fonction du mode articulatoire, facteur à 3 niveaux : les occlusives /p, b, t, d, k, g/, les fricatives /f, v, s, z, ʒ, ʃ/ et les consonnes sonnantes (ou vocaliques) /m, n, l, R/. La figure 27 nous montre que les consonnes occlusives présentent des scores plus élevés que les consonnes fricatives et sonnantes. Les occlusives sont donc plus altérées chez nos patients, et ceci est significatif ($p < 2.2^{-16}$).

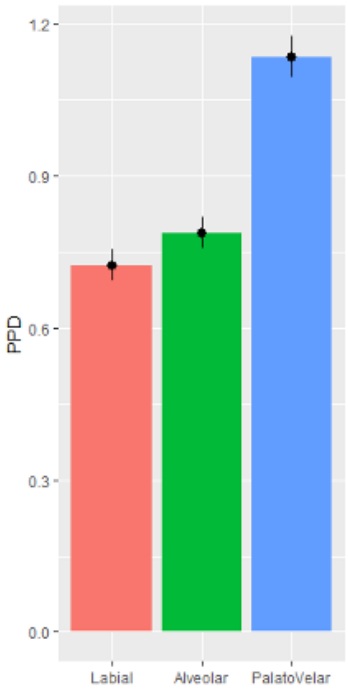

Figure 26 : Scores PPD locaux pour les consonnes en fonction du lieu d'articulation : labial, alvéolaire ou palato-vélaire.

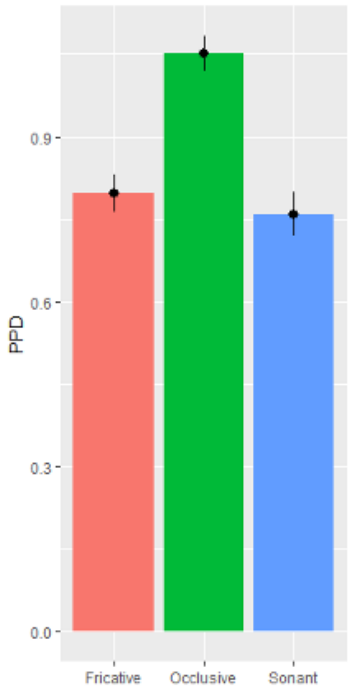

Figure 27 : Scores PPD locaux des consonnes en fonction de leur mode d'articulation : occlusif, fricatif, et enfin les consonnes sonnantes.

V. DISCUSSION

Le but de notre travail était d'analyser l'impact que peuvent avoir la taille d'une tumeur et sa localisation sur l'intelligibilité de la parole après traitement d'un cancer de la cavité buccale ou de l'oropharynx, et ce en utilisant un nouvel outil d'évaluation de l'intelligibilité, la tâche de décodage acoustico-phonétique, afin de pallier le manque de connaissances quant aux conséquences d'un traitement d'un cancer ORL sur la parole mais aussi au manque d'outil pour une évaluation fiable et objective de l'intelligibilité.

Nos résultats sur les scores PPD globaux ont montré que l'altération de l'intelligibilité ne dépend pas de la région atteinte par la tumeur (cavité buccale vs oropharynx). Ceci peut être dû soit à un manque de données, mais notre corpus comportait quand même 85 patients (33 pour la cavité buccale et 52 pour l'oropharynx), soit car il n'y a pas réellement d'effet de la région atteinte sur l'intelligibilité, mais simplement une tendance. Balaguer et al. (2019) montraient que les patients avec un cancer de la cavité buccale présentaient une intelligibilité plus altérée et une altération du signal de parole plus importante que les patients avec un cancer de l'oropharynx. Borggreven et al. (2005) constataient une altération de l'intelligibilité plus importante pour les cavités orales avant traitement, mais au contraire de meilleurs résultats pour les cavités orales après traitement, sans que cela soit significatif. Nous retrouvons donc des résultats différents à travers les différentes études quant à l'effet de la localisation d'une tumeur sur l'intelligibilité. Un quelconque effet de la localisation n'est donc peut-être pas si certain. Mais nous pouvons aussi discuter de l'intérêt de cette dichotomie cavité buccale / oropharynx en linguistique. En effet, dans le cadre de la clinique et du domaine de l'ORL cette répartition est parfaitement adaptée, il s'agit d'une description anatomique qui permet de classer tel ou tel organe dans l'une des deux régions. En revanche, parler de cavité buccale ou d'oropharynx lorsqu'il s'agit d'intelligibilité de la parole semble plus discutable, nos résultats n'ayant pas donné de résultats probants en fonction du facteur région pour les scores globaux, au contraire des scores locaux qui semblent influencés par les localisations précises des tumeurs. Cette dichotomie est peut-être trop grossière pour discuter d'intelligibilité. Il est donc possible que certains outils de classification utilisés en ORL ne soient pas utilisables dans ce domaine différent qu'est la linguistique.

Avec les scores PPD globaux, nous avons également constaté un effet largement significatif de la taille de la tumeur sur les scores d'intelligibilité. Nous avons initialement regardé les scores en fonction des quatre différentes tailles, T1 à T4. Seuls les scores des tumeurs classées T4 étaient significativement plus hauts que les scores des tumeurs classées T2. L'absence de différence significative entre les autres groupes de taille ne nous convenait

pas, puisque si une tumeur T4 altère plus l'intelligibilité qu'une tumeur T2, elle doit logiquement l'altérer encore plus qu'une tumeur T1. Mais la solution nous paraissait claire lorsque nous regardions le corpus : nous comptions 36 tumeurs T2 et 27 tumeurs T4 pour seulement 10 tumeurs T1 et 12 tumeurs T2. Il y avait donc une disparité d'effectifs entre les différentes tailles de tumeur. En choisissant de regrouper les petites tumeurs T1 et T2, et les grosses tumeurs T3 et T4, nous créons ainsi deux groupes relativement équilibrés avec respectivement 46 et 39 patients. Les nouvelles analyses effectuées sur ces deux groupes ont permis de faire ressortir un effet significatif de la taille de la tumeur sur l'intelligibilité. Plus une tumeur est grosse, plus l'altération de l'intelligibilité est importante. Ceci est parfaitement compatible avec d'autres études antérieures, par exemple Balaguer et al. (2019) qui constataient une plus grosse altération de l'intelligibilité pour les grosses tumeurs T4, Borggreven et al. (2005), Thomas et al. (2008) ou encore Manrique et al. (2016), ces derniers ayant la même constatation mais pour les tumeurs linguales en particulier. Borggreven avait d'ailleurs regroupé de la même manière les grosses tumeurs T3-T4 pour les confronter aux petites tumeurs T2 (pas de tumeur T1 dans son étude), pour des raisons d'effectifs également. De tels résultats sont explicables par plusieurs faits. D'une part par l'atteinte directe de la mobilité des structures adjacentes à la tumeur impliquées dans la parole. Plus une tumeur est grosse, plus elle va atteindre les structures autour et réduire leur mobilité. D'autre part, la taille d'une tumeur est associée à la lourdeur du traitement. En effet, les petites tumeurs T1-T2 auront tendance à être traitées par monothérapie (chirurgie ou radiothérapie), contrairement aux grosses tumeurs T3-T4 qui souvent bénéficieront de deux traitements associés voire trois (chirurgie et/ou radiothérapie et/ou chimiothérapie). Les effets secondaires des traitements s'accumulent donc. De même, une grosse tumeur T3-T4 si elle est traitée chirurgicalement impliquera une plus grosse perte de substance (dans la cavité orale ou l'oropharynx) qu'une petite tumeur T1-T2. Cette notion de volume de résection tumorale est d'ailleurs retrouvée dans plusieurs études (Kreeft et al., 2009 ; Borggreven et al., 2005), celui-ci étant inversement corrélé à l'intelligibilité. Pauloski et al. (1998) établissent en outre une corrélation entre le pourcentage de tissu réséqué (de langue, de plancher buccal, …) et le pourcentage de phonèmes consonantiques reconnus.

Concernant les scores PPD locaux, nos résultats ont montré que les scores d'intelligibilité étaient significativement plus hauts (donc plus altérés) pour les clusters de consonnes (type CC), puis pour les voyelles puis enfin pour les consonnes. Nos patients étaient donc moins intelligibles lors de la prononciation des clusters de consonnes puis des voyelles, alors que les consonnes simples semblent être plus aisément réalisées. Ceci est intéressant puisque cela nous permet d'avancer quant à la question posée par le projet RUGBI, à savoir

quelles sont les unités linguistiques pertinentes pour améliorer l'évaluation de l'intelligibilité dans la parole pathologique ? La piste des clusters de consonnes est donc à privilégier, puisque les patients traités pour un cancer de la cavité buccale ou de l'oropharynx semblent être plus en difficulté pour la production de ce type de phonèmes, ces derniers impliquant des gestes articulatoires plus complexes. Cela serait notamment pertinent dans le cadre de réduction de liste, explorée par Marczyk et al. (2020) : une liste d'items équivalente à 30% de la liste originelle, si elle est composée non pas arbitrairement mais en conservant les segments difficilement réalisables par les patients (comme les clusters), permettrait un gain de temps tout en conservant une évaluation fiable et robuste.

Parmi les voyelles, ce sont les voyelles nasales qui étaient le plus altérées chez nos patients, avec des scores locaux plus de deux fois supérieurs aux scores locaux des voyelles orales (respectivement 1.98 et 0.81, $p < 2.2^{-16}$), et un effet d de Cohen moyen. La nasalité est donc un problème pour ces patients. Lorsque nous recoupons cette donnée avec l'effet de la région de la tumeur, on s'aperçoit que cet effet est majoré : les voyelles nasales sont encore plus altérées pour les tumeurs de l'oropharynx, et les voyelles orales pour les tumeurs buccales. Cela est compatible avec la littérature, puisque Borggreven et al. (2005) retrouvaient une altération plus importante de la nasalité chez les patients atteints de cancers de l'oropharynx (amygdale et palais mou) que chez les patients avec un cancer de la cavité buccale. La région a donc ici un effet. C'était également le cas pour Kreeft et al. (2009). Rieger et al. (2008) montraient que la nasalité était dépendante du volume de palais mou réséqué. Nous proposons plusieurs explications à cette altération de la nasalité, prédominante pour les tumeurs de l'oropharynx. D'une part, les traitements ciblant l'oropharynx, que ce soit la chirurgie ou la radiothérapie, vont atteindre quasi systématiquement le voile du palais, responsable de la nasalité via son mouvement de fermeture ou d'ouverture entre les cavités buccale et nasale. Ces traitements entraînent une fibrose voire une exérèse du voile, ce dernier devenant ainsi moins mobile. L'évaluation de la nasalité semble donc justifiée dans le suivi des patients traités pour une tumeur de la cavité buccale ou de l'oropharynx.

Une fois les voyelles nasales traitées, nous nous sommes intéressés aux voyelles orales seules, d'abord par rapport à leur réalisation sur l'axe antéro-postérieur. Nous avons constaté une altération significativement plus importante des voyelles postérieures /Ô et u/ par rapport aux voyelles antérieures /E i y/. Cette altération est majorée lorsqu'il s'agit d'une tumeur de la cavité buccale. Nous avons également analysé les voyelles orales selon leur degré d'aperture, et là encore nous obtenions une différence significative puisque les voyelles hautes /i, u, y/ présentaient des scores PPD locaux plus élevés que les voyelles moyennes /E, Ô/ et la voyelle basse /a/. Les voyelles orales postérieures et hautes sont donc les plus altérées. Si l'on reprend

le triangle vocalique ci-dessous (Fig.28), qui représente le degré d'aperture des voyelles et leur lieu de réalisation sur l'axe antéro-postérieur, il est plus aisé de comprendre le rôle essentiel qu'a la langue dans ces deux traits. En effet ce sont ses mouvements d'avant en arrière et de bas en haut (avec également une participation de la mandibule pour l'aperture) qui permettent de varier de lieu d'articulation et donc de produire différentes voyelles (Fig.29). Pour ces dernières, ce sont des mouvements plus ou moins grossiers de la langue qui vont entrer en jeu, puisque l'on a un déplacement de la masse linguale dans sa totalité grâce notamment aux muscles extrinsèques. Lorsque la langue est atteinte, on comprend que la réalisation des voyelles soit plus compliquée, et notamment des voyelles orales postérieures et hautes puisque ce sont elles qui nécessitent des plus gros déplacements. Si l'on considère que la langue est constituée de la langue mobile, dans la cavité orale, et de la base de langue, dans l'oropharynx, on peut se demander si la dichotomie de région cavité buccale vs oropharynx est réellement adaptée, puisque la langue est présente dans les 2 régions. En revanche, la question de l'impact de la localisation précise d'une tumeur sur l'intelligibilité, ou d'un ensemble d'articulateurs, est pertinente. Il serait donc plus avisé, non pas de réfléchir en termes de région cavité buccale vs oropharynx (dichotomie anatomique adaptée à la clinique, mais finalement peut-être pas à la linguistique), mais plutôt en termes de localisation précise ou d'ensemble d'articulateurs.

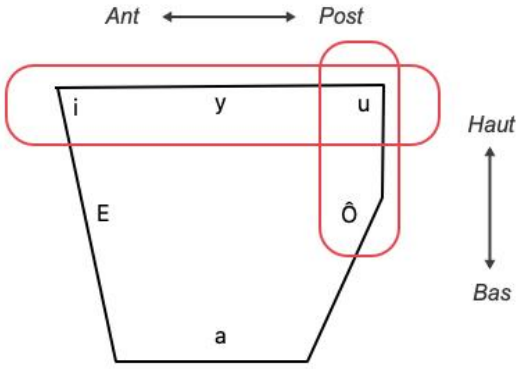

Figure 28 : Triangle vocalique représentant schématiquement les voyelles orales en fonction de leur lieu d'articulation (position de la langue selon l'axe antéro-postérieur) et de l'aperture. Les zones rouges symbolisent les regroupements des voyelles hautes et des voyelles postérieures, c'est-à-dire les voyelles pour lesquelles la langue est haute et/ou postérieure.

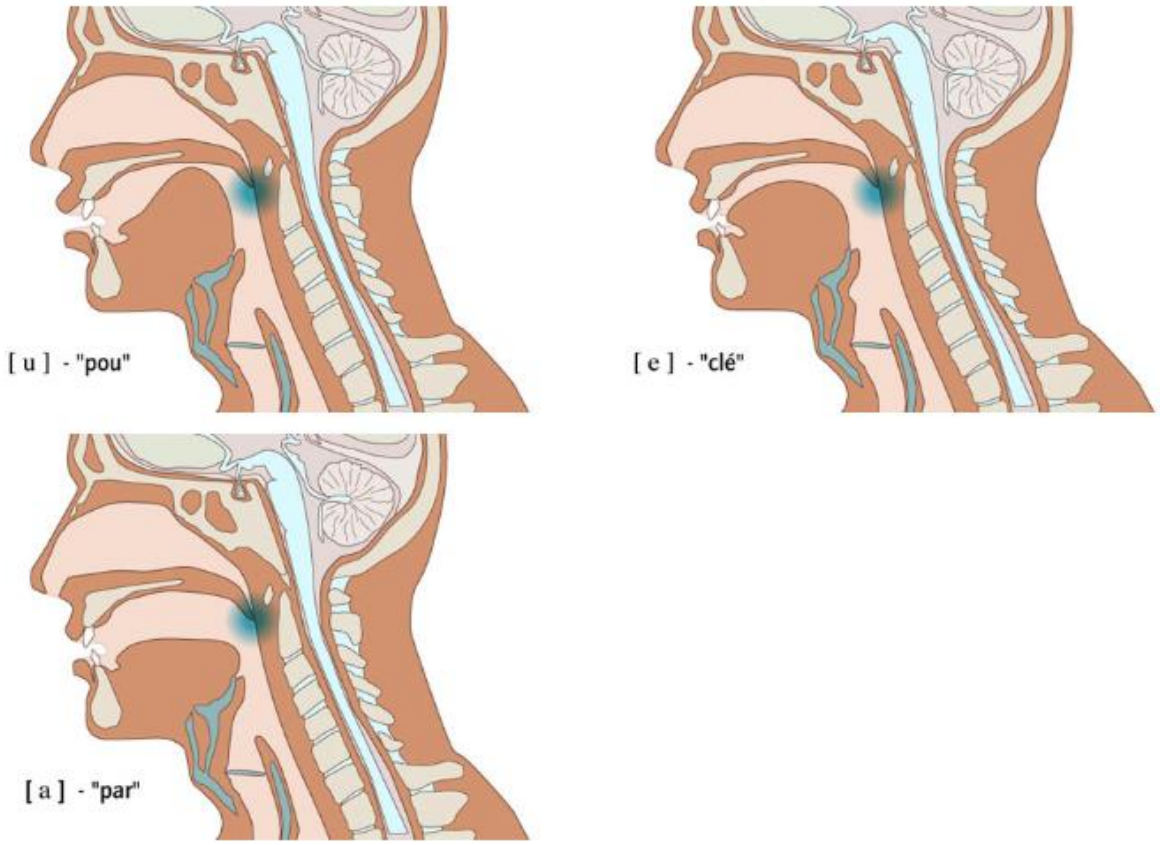

Figure 29 : Coupes sagittales du conduit vocal, lors de l'articulation des voyelles /u/ (en haut à gauche), /e/ (en haut à droite) et /a/ (en bas). On peut apercevoir les variations de positions de la langue, qui pour le /u/ est haute et en arrière, légèrement plus basse et en avant pour /e/, et basse et moyennement en avant pour /a/. C'est toute la masse linguale ici qui se mobilise, grâce notamment aux muscles extrinsèques.

Enfin, les scores PPD locaux nous ont permis d'analyser les consonnes, en fonction de leur lieu d'articulation et de leur mode articulatoire. L'intelligibilité était significativement plus altérée pour les consonnes palato-vélaires /ʒ, ʃ, k, g, R/ donc très postérieures, que pour les consonnes alvéolaires /t, d, s, z, n, l/ ou labiales /p, b, f, v, m/. Nous retrouvons encore cette notion de postériorité associée à l'altération de l'intelligibilité, que nous avions pour les voyelles. Borggreven et al. (2005), en se portant sur le taux d'erreurs consonantiques en fonction des localisations de tumeurs, s'étaient aperçus que les consonnes /k, s, d, t/ étaient les plus difficiles à réaliser pour leurs patients, avec un taux d'erreurs majeurs pour les tumeurs amygdaliennes (19%) et faible pour les tumeurs de la base de langue (11%), autrement dit deux localisations oropharyngées. Cette étude ne mettait donc pas en évidence un lien franc entre la région d'une tumeur et la prononciation des consonnes. Au vu de nos résultats, la notion de région (cavité buccale vs oropharynx) ne semble pas non plus jouer sur les consonnes, mais bien encore une fois la langue (base de langue), ainsi que le voile du palais, qui sont les deux organes les plus impliqués dans la réalisation des consonnes palatales. Cela constitue un argument supplémentaire pour évoquer la langue dans sa globalité, à savoir la langue mobile et la base de langue, comme un organe majeur dans l'intelligibilité de la parole, cette dernière étant

fortement dégradée lorsque la langue est touchée par la tumeur. Ceci n'est pas étonnant, puisque la réalisation des consonnes implique des mouvements fins et précis de la musculature intrinsèque de la langue, à l'inverse des voyelles réalisées à l'aide de mouvements plus grossiers de la langue, ce que nous avons évoqué plus haut. Pour terminer sur les consonnes, nous avons vu que les consonnes occlusives sont plus altérées chez nos patients que les consonnes fricatives ou sonnantes. Ce sont donc les consonnes impliquant une fermeture complète du tractus vocal qui sont donc les plus atteintes. Cela nous paraît logique, puisqu'une constriction complète du tractus nécessite plus d'effort et d'amplitude de mouvement, rendant ainsi ce mouvement plus difficile à réaliser qu'une simple constriction partielle. Une des explications à cela est notamment la fibrose et la rigidité post-thérapeutiques des tissus du tractus vocal, quand ceux-ci n'ont pas été retirés au cours d'une chirurgie. Ils sont ainsi beaucoup plus limités dans leurs mouvements.

A l'issue de ce travail, nous pouvons affirmer que l'intelligibilité de la parole est altérée au cours d'un cancer de la cavité buccale ou de l'oropharynx. La taille de la tumeur est fortement associée à cette altération. Ainsi lorsque nous sommes confrontées à une volumineuse tumeur, nous pouvons prédire que l'intelligibilité du patient sera bien plus altérée que s'il avait présenté une tumeur de petite taille. Les régions cavité buccale et oropharynx au sens anatomo-clinique ne semblent en revanche pas intervenir, du moins au niveau global, mais plutôt au niveau segmental, avec probablement la localisation linguale précisément. Ainsi nos intuitions de base, qui étaient notamment qu'un cancer antérieur (de la cavité buccale) allait altérer préférentiellement les phonèmes antérieurs, et un cancer postérieur allait altérer plutôt les phonèmes postérieurs, n'étaient pas si vraies. La réalité semble plus complexe, et il sera intéressant dans des projets à venir de se porter plutôt sur les localisations précises des tumeurs (langue, mandibule, etc…) voire même d'analyser les résultats en fonction de groupes d'articulateurs, et notamment centrés autour de la langue. Nous avons également vu que les patients sont majoritairement altérés sur la prononciation des clusters de consonnes, des consonnes palatales et des voyelles nasales, orales postérieures et hautes. C'est l'évaluation de ces phonèmes en particulier qui semble donc pertinente pour l'évaluation de l'intelligibilité chez les patients traités pour un cancer de la cavité buccale ou de l'oropharynx. Ces constatations ont pu être possibles grâce à l'utilisation de la tâche de décodage acoustico-phonétique, qui se révèle être un outil fiable et objectif dans le domaine clinique, et qui à terme pourrait être utilisé par le clinicien (médecin, orthophoniste). Ce travail ouvre en outre plusieurs perspectives. Seuls les phénomènes de substitution ont été analysés dans ce travail. Un futur

projet pourra se porter sur les phénomènes d'insertion et d'élision, qui pourront révéler d'autres profils d'altération. De plus, il serait intéressant d'étudier la relation entre le niveau d'altération des phonèmes et la diversité des réponses proposées par les évaluateurs, l'hypothèse étant que plus un phonème est altéré, plus les auditeurs proposent des réponses diverses. Il semblerait en effet que nos consonnes palatales /k/ et /g/ par exemple, soient associées à un grand nombre de réponses différentes (cinq ou six types de réponses majoritaires), à l'inverse de la consonne /m/ qui est peu altérée et qui présente seulement deux réponses majoritaires possibles (/m/ et /n/). Enfin, il serait très informateur de comparer les résultats de nos patients avec les résultats des sujets contrôles, afin de savoir quelles altérations relèvent de la pathologie tumorale et lesquelles relèvent de la parole tout simplement. C'est un travail que Marie Rebourg, membre du projet RUGBI, a déjà initié. Il semblerait par exemple que les voyelles nasales soient également altérées chez les sujets contrôles, la nasalité est peut-être donc très fragile de base, et plus susceptible d'être altérée que d'autres segments.

REFERENCES

Astésano C, Balaguer M, Farinas J, Fredouille C, Ghio A, et al. Carcinologic Speech Severity Index Project: A Database of Speech Disorder Productions to Assess Quality of Life Related to Speech After Cancer. Eleventh International Conference on Language Resources and Evaluation Conference (LREC 2018), May 2018, Miyazaki, Japan. hal-01770168

Balaguer M, Boisguerin A, Galtier A, Gaillard N, Puech M, Woisard V. Factors influencing intelligibility and severity of chronic speech disorders of patients treated for oral or oropharyngeal cancer. Eur Arch Otorhinolaryngol 2019;276(6):1767-1774

Balaguer M, Boisguerin A, Galtier A, Gaillard N, Puech M, Woisard. Assessment of impairment of intelligibility and of speech signal after oral cavity and oropharynx cancer. Eur Ann Otorhinolaryngol Head Neck Dis 2019;136(5):355-359

Blanc E, Giusti L. (2014). Évaluation de l'intelligibilité de la parole dans les dysarthries. Adaptation en français de la version révisée du Frenchay Dysarthria Assessment (FDA-2). Mémoire de fin d'études pour l'obtention du Certificat de Capacité en Orthophonie. Aix-Marseille Université, école d'orthophonie.

Borggreven PA, Verdonck-de Leeuw IM, Muller MJ, et al. Quality of life and functional status in patients with cancer of the oral cavity and oropharynx: pretreatment values of a prospective study. Eur Arch Otorhinolaryngol 2007;264(6):651–7

Borggreven PA, Verdonck-de Leeuw IM, Langendjik JA, et al. Speech outcome after surgical treatment for oral and oropharyngeal cancer: a longitudinal assessment of patients reconstructed by a microvascular flap. Head & Neck 2005;27(9):785-793

Chomsky N., Halle M (1968), The Sound Pattern of English. New York: Harper & Row

Debruijn M, Ten Bosch L, Kuik DJ, Quené H, Langendijk JA, Leemans CR, Verdonck-de Leeuw I. Objective Acoustic-Phonetic Speech Analysis in Patients Treated for Oral or Oropharyngeal Cancer. Folia Phoniatr Logop 2009;61(3):180-7

Ghio A, Giusti L, Blanc E, Pinto P, Lalain M, et al. Quels tests d'intelligibilité pour évaluer les troubles de production de la parole ? Journées d'Etude sur la Parole, 2016, Paris, France : 589-596. hal-01372037

Ghio A, Lalain M, Rebourg M, Fredouille C, Woisard V. Évaluation de l'intelligibilité de patients avec traitement du cancer des cavités orales et pharyngales. 6e conférence conjointe Journées d'Études sur la Parole (JEP, 33e édition), Traitement Automatique des Langues Naturelles (TALN, 27e édition), Rencontre des Étudiants Chercheurs en Informatique pour le Traitement Automatique des Langues (RÉCITAL, 22e édition). Volume 1 : Journées d'Études sur la Parole, 2020, Nancy, France :271-279. hal-02798549v3

Ghio A, Lalain M, Giusti L, Pouchoulin G, Danièle Robert D, et al. Une mesure d'intelligibilité par décodage acoustico-phonétique de pseudo-mots dans le cas de parole atypique. XXXIIe Journées d'Etudes sur la Parole, LPL, 2018, Aix-en-Provence, France :285-293, 10.21437/jep.2018-33. hal-01770161v2

Jakobson R, Fant G, Halle M (1951). Preliminaries to speech analysis. MIT Press, Cambridge.

Keintz CK, Bunton K, Hoit JD. Influence of Visual Information on the Intelligibility of Dysarthric Speech. American Journal of Speech-Language Pathology 2007;16(3):222-234

Kreeft AM, Van Der Molen L, Hilgers FJ, Balm AJ. Speech and swallowing after surgical treatment of advanced oral and oropharyngeal carcinoma: a systematic review of the literature. Eur Arch Otorhinolaryngol 2009;266(11):1687-1698

Lalain M, Ghio A, Fredouille C, Rebourg M, Belleville L, et al.. Mots ou pseudo-mots ? Quel impact sur l'évaluation perceptive de l'intelligibilité ? Journées de Phonétique Clinique, Mai 2019, Mons, Belgique. hal-02482566

Lalain M, Ghio A, Giusti L, Robert D, Fredouille C, Woisard V. Design and Development of a Speech Intelligibility Test Based on Pseudowords in French: Why and How? J Speech Lang Hear Res 2020;63(7):2070-2083

Marczyk A, Ghio A, Lalain M, Rebourg M, Fredouille C, Woisard V. Have a Cake and Eat it Too : Assessing Discrimination Performance of an Intelligibility Index Obtained from a Reduced Sample Size. 12th Conference on Language Resources and Evaluation (LREC 2020), ELRA, May 2020, Marseille, France. pp.1784-1788. hal-02611678

McKinstry A, Perry A. Evaluation of speech in people with head and neck cancer: a pilot study. Int J Lang Commun Disord 2003;38(1):31–46

Pauloski BR, Logemann JA, Colangelo LA, Rademaler AW, McConnel FM, Heiser MA, Cardinale S, Shedd D, Stein D, Beery Q, Myers E, Lewin J, Haxer M, Esclamado R. Surgical Variables Affecting Speech in Treated Patients With Oral and Oropharyngeal Cancer. Laryngoscope 1998;108(6):908-16

Rebourg M, Lalain M (2017). Mesure de l'intelligibilité : la tâche de décodage acoustico-phonétique dans le développement d'outils de mesure. Mémoire de fin d'année, pour l'obtention du master 1 Sciences du langage : linguistique expérimentale. Aix-Marseille Université, Laboratoire Parole et Langage.

Rebourg M, Lalain M, Ghio A, Fredouille C, Fakhry N, Woisard V. Evaluer l'intelligibilité, mots ou pseudo-mots ? Comparaison entre deux groupes d'auditeurs. 6e conférence conjointe Journées d'Études sur la Parole (JEP, 33e édition), Traitement Automatique des Langues Naturelles (TALN, 27e édition), Rencontre des Étudiants Chercheurs en Informatique pour le Traitement Automatique des Langues (RÉCITAL, 22e édition). Volume 1 : Journées d'Études sur la Parole, 2020, Nancy, France. pp.543-551. hal-02798584v3

Thomas L, Jones TM, Tandon S, Carding P, Lowe D, Rogers S. Speech and voice outcomes in oropharyngeal cancer and evaluation of the University of Washington Quality of Life speech domain. Clin Otolaryngol 2009;34(1):34-42

Woisard V, Astésano C, Balaguer M, Farinas J, Fredouille C, et al. C2SI corpus: a database of speech disorder productions to assess intelligibility and quality of life in head and neck cancers. *Language Resources and Evaluation*, Springer Verlag, 2020. hal-02921918

Yorkston KM, Dowden PA, Beukelman DR. Intelligibility measurement as a tool in the clinical management of dysarthric speakers», in Intelligibility in speech Disorders : Theory,

measurement and management. Raymond D. Kent. Madison, Wisconsin: John Benjamins Publishing Company, 1992:265-286