

Liste des Figures

Figure 1. Distribution de la vitesse de parole en fonction du groupe (dans chaque situation d'analyse différente).	8
Figure 2. Durée d'émission de voyelle en pourcentage (V%) en fonction du quotient de variabilité de la durée des voyelles (ΔV), avec les barres d'erreur représentant les écarts-types. Gauche : <i>Jean mène un journal</i> . Droite : <i>J'emmène un journal</i>	10
Figure 3. Répartition des scores de chaque sous-partie du VHI selon le Groupe (VO ou VTO) dans différents degrés de sévérité.....	11
Figure 4 : Corrélation entre score Total du VHI et la vitesse de parole en lecture. Nuage de points et courbe de régression linéaire.....	12

Liste des tableaux

Tableau 1 : Corrélations entre dimensions de qualité de vie (EORTC QLQ H&N 35) et sous-scores du VHI	12
--	----

Liste des abréviations

LT = laryngectomie totale

SSO = sphincter supérieur de l'œsophage

VO = voix oro-œsophagienne

VTO = voix trachéo-œsophagienne

T = témoin

ORL = oto-rhino-laryngologie

CV = consonne-voyelle

CCV = consonne-consonne-voyelle

VHI = *Voice Handicap Index*

EORTC QLQ H&N 35 = (*European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire Head and Neck Module*)

%V = Vperc, soit la durée des intervalles vocaliques par rapport à tout l'énoncé

ΔV = Vdev soit la variabilité de la durée des intervalles vocaliques

Résumé

Cette étude dresse un état des lieux de la parole après laryngectomie totale (LT) en se focalisant sur les paramètres temporels globaux (d'un énoncé) et locaux (à partir de segments de parole). Un corpus composé d'items issus de la littérature a permis d'évaluer les différents paramètres temporels dans trois situations (lecture, discours spontané et diadococinésies). Il a été soumis à une population composée de 12 sujets utilisant la voix oro-œsophagienne (VO), 17 la voix trachéo-œsophagienne (VT) et de 12 sujets témoins (T). Les résultats ont montré que la vitesse de parole des patients, en particulier de ceux utilisant la VO, était ralentie, dans toutes les situations. La fréquence des pauses et le nombre d'expirations élevés démontrent aussi un impact réel de la LT sur la rapidité et la fluidité de l'élocution. Des mesures plus précises autour des voyelles ont été réalisées : le groupe VO se distingue des groupes VTO et T par une plus grande variabilité des voyelles, et le groupe T a des résultats plus contrastés entre les deux phrases enregistrées alors qu'ils sont plus stables chez les groupes VO et VTO. Les perturbations de la parole peuvent troubler la qualité de vie et sont en lien avec la voix. Des corrélations positives significatives ont été observées entre des items de qualité de vie et de handicap vocal ; de plus, le handicap vocal des patients LT serait plus élevé lorsqu'ils présentent une vitesse de lecture basse et inversement.

Mots clefs : laryngectomie totale, rythme, prosodie, phonétique

Abstract

The present study is a state of the art of speech ability after total laryngectomy. It focuses on global (of an utterance) and local (speech segments) temporal parameters of elocution. A text corpus based on literature extracts enabled to evaluate various temporal parameters through three categories (reading, spontaneous speech and diadokokinesis). The corpus was submitted to 12 subjects using oesophageal voice, 17 subjects using tracheoesophageal voice and 12 control subjects. Results show that patients' speech velocity was reduced for the three categories of communication, in particular for the patients with oesophageal voice. Pauses frequency as well as the expiration number also show a significant impact of total laryngectomy on elocution velocity and fluidity. The vowels measurements were performed with more precision: the oesophageal group presented longer vowel durations than other groups. In addition, the vowel times were observed to differ more (between the recorded pairs of sentences) for the control group than other groups. Speech malfunction might trouble life quality and are linked to voice quality and voice handicap. Significant correlations were observed between some aspects of life quality and voice handicap; moreover, laryngectomised subjects were shown to feel a strong voice handicap when presenting a slow read velocity and reverse.

Keywords: total laryngectomy, rhythm, prosody, phonetics

Introduction

La laryngectomie totale (LT) est une chirurgie qui prive le patient de ses plis vocaux et donc de sa voix naturelle. Lors de la respiration, l'air inspiratoire et expiratoire ne traverse plus le pharynx ni les cavités buccale et nasale. Il est acheminé entre les poumons et le trachéostome, à la base du cou. Afin que les patients puissent continuer à s'exprimer, deux voix de substitution sont possibles. La « voix oro-œsophagienne » (VO), utilise l'air buccal, « injecté » par pressions sous le sphincter supérieur de l'œsophage (SSO) qui joue le rôle de vibreur pour permettre l'émission de la voix (Giovanni et Robert, 2010). La « voix trachéo-œsophagienne » (VTO) est produite à partir de l'air expiratoire, grâce à un implant reliant la trachée à l'œsophage, en regard du trachéostome. La structure vibrante est identique à celle de la VO (le SSO). Le timbre de voix pour ces deux types de réhabilitation est donc analogue.

La perte du vibreur laryngé perturbe la prosodie. L'effet est plus important pour les patients VO que les VTO avec une grande différence entre les deux (Van Rossum, 2008). La prosodie des LT présente une intonation réduite (Giovanni et Robert, 2010), une intensité faible (Štajner-Katušić *et al.*, 2006) ainsi qu'un débit de parole perturbé, en particulier pour la VO (Le Huche, Allali et Hilgers, 2015 ; Crevier-Buchman, Vaissière, Maeda et Brasnu, 2002). La littérature s'accorde pour affirmer l'influence considérable de la LT sur le rythme de la parole et la différence notable entre la VO (plus impactée) et la VTO (moins impactée grâce à la conservation du lien pneumo-phonique). En français, le rythme joue un rôle primordial dans la compréhension de la parole (Caelen-Haumont, 2016 ; Duez, Legou et Viallet, 2009) : (1) au niveau de la phrase, par ses contraintes syntaxiques et grammaticales (Canault, 2017) ; (2) au niveau syllabique, avec une régularité temporelle caractéristique (Fitch, 2019 ; Pillot-Loiseau et Xie, 2018) ; (3) au niveau du phonème, où les notions de durée et de variabilité contribuent à formuler des unités de sens (Pillot-Loiseau et Xie, 2018 ; Mairano et Romano, 2011). Le manque d'intelligibilité des patients LT utilisant la VO est donc directement lié à un trouble de la fluidité du discours et de ses aspects rythmiques (Motta, 2001).

Dans ce contexte, nous nous interrogeons sur les conséquences de la LT sur le rythme de la parole. Un protocole comprenant des épreuves innovantes nous permettra : d'identifier des paramètres rythmiques locaux - jamais encore soumis à une population LT – et locaux, impactés chez ces patients, de caractériser l'impact de la LT en s'attachant au lien avec la qualité de vie et plus particulièrement à l'impact d'une perturbation de la parole sur le handicap vocal ressenti.

Matériel et Méthodes

Nous avons recruté des patients ayant subi une LT selon des critères d'inclusion et d'exclusion prédéfinis. L'âge des sujets devait être compris entre 25 et 85 ans. Tous les patients devaient avoir subi une laryngectomie totale et utiliser la VTO ou la VO. Nous demandions que le sujet ait bénéficié d'une rééducation orthophonique après son opération. Un délai de 6 mois post-opératoire et post traitement radio/chimio-thérapeutique a été fixé. Le dernier critère d'inclusion consistait à savoir correctement lire le français. Nous avons ensuite choisi les critères d'exclusion suivants pour les deux groupes de patients : l'analphabétisme ou l'illettrisme, toute pathologie neurologique ou langagière et la surdité. Les critères d'inclusion du groupe témoin (T) concernent aussi l'âge (25 à 85 ans) ainsi qu'une bonne maîtrise de la lecture en français. Les critères d'exclusion du groupe témoin sont les mêmes que pour les patients soient l'analphabétisme ou l'illettrisme, et toute surdité ou pathologie neurologique ou langagière.

L'acquisition des données s'est effectuée grâce à la contribution d'orthophonistes des secteurs hospitaliers ORL (Centre Hospitalier Intercommunal de Créteil, Institut Gustave Roussy, Institut Universitaire du Cancer de Toulouse et les Hôpitaux René-Dubos, Pitié Salpêtrière, George Pompidou et Tenon) qui nous ont permis de rencontrer les sujets soit à l'hôpital, soit à leur domicile. Nous avons aussi bénéficié de la participation de membres actifs de l'AFLMV (Association Française des Laryngectomisés et Mutilés de la voix) d'Ile de France et d'Occitanie. Les sujets que nous avons pu effectivement rencontrer sont au nombre de 41. Deux sujets des groupes de patients ont dû annuler l'entretien de passation pour cause de récurrence de leur cancer. Deux sujets n'ont pas pu être inclus à cause d'un accent étranger très prononcé. Un sujet n'a pas pu participer pour cause d'analphabétisme et un autre pour cause de dyslexie. Deux patients ont refusé de participer. Les femmes étant peu nombreuses dans la population LT, nous avons recruté une grande majorité d'hommes pour les groupes de patients. Pour les groupes témoins, nous avons décidé le plus possible de ne pas recruter de sujet de moins de 50 ans, de sorte que l'âge ne soit pas un facteur de confusion trop important lors du traitement de nos données. Ainsi, de nombreux participants potentiels (surtout des hommes) ont été écartés car ayant déjà eu un ou plusieurs AVC.

Pour les enregistrements, nous avons utilisé un micro-casque *AKG*, situé à deux doigts de la bouche du sujet, une carte son *Roland* et le logiciel Audacity (version 2.1.0.0). Une version papier du protocole est mise en possession du patient lors de la passation ainsi que les grilles des questionnaires et la fiche de consentement qu'il doit remplir. Pour le traitement des données, nous avons utilisé les logiciels Praat (version 6.0.33) et Correlatore

(Mairano et Romano, 2011). Pour le traitement statistique, nous avons utilisé JMP (version 14.1.0) et Excel.

Notre protocole de passation dure trente minutes. Il s'organise comme suit : une partie orale (enregistrement) de vingt minutes puis une partie écrite de dix minutes. L'enregistrement du patient a été réalisé à partir d'un corpus qui regroupe différents types d'épreuves (voyelle tenue, phrases sur plusieurs intonations et avec différents rythmes, voir plus bas), nécessaires à une récolte la plus exhaustive possible d'échantillons de parole, à analyser d'un point de vue rythmique. Une fois l'enregistrement terminé, le patient a rempli à l'écrit les questionnaires de qualité de vie et de handicap vocal.

Le corpus débutait par un Temps maximal de Phonation (TMP) sur la voyelle /a/. Puis, une première partie a réuni différentes épreuves de lecture ; en premier lieu, celle de la phrase « Mélanie vend du lilas », issue du protocole *MonPage* (Pernon *et al.*, 2020), dans laquelle chaque phonème est voisé. Nous avons ensuite proposé la lecture de cette phrase avec trois intonations différentes imposées (affirmation, interrogation puis exclamation), puis celle de six couples de phrases ambiguës (un couple comporte deux phrases dont on distingue le sens uniquement par une modification prosodique), proposées par Vaissière (2015) avec :

- 1 : Elle manque à mon amie / Elle manqua mon amie
- 2 : Cet homme est énormément bête / Cet homme est énorme et m'embête
- 3 : Il aime à en mourir / Il aima en mourir,
- 4 : Je suis inquiet de la varice de ma mère / Je suis inquiet de l'avarice de ma mère
- 5 : Jean mène un journal / J'emmène un journal
- 6 : Le rapace la noie / Le rat passe la noix

Cette partie du corpus, orientée sur la lecture, s'est achevée par une lecture de récit. Elle permet de tester la parole du patient avec la contrainte de la longueur ajoutée. Nous avons proposé la lecture du texte *La bise et le soleil*, récit fréquemment retrouvé dans les études en phonétique française, et qui comporte tous les phonèmes de la langue. La deuxième partie du corpus comportait une épreuve de diadococinésies. Une série de diadococinésies a donc été proposée sur des syllabes ciblées ayant volontairement un contexte consonne-voyelle (CV) rapproché, comme proposé dans le protocole *Monpage* (Pernon *et al.*, 2020). Nous avons donc obtenu six séries de trois syllabes identiques de différents niveaux de complexité, avec 3 syllabes de type consonne-voyelle (CV) : « ba ba ba », « de de de », « go go go », et trois syllabes de type consonne-consonne-voyelle (CCV) : « cla cla cla », « tra tra tra », « cla tra ». Le patient devait répéter le plus rapidement et le plus vite possible chaque série de trois syllabes pendant cinq secondes (que nous avons chronométrées). Pour finir, nous avons demandé au patient de nous raconter ce qu'il avait

fait avant de venir nous rencontrer. Cela nous a permis de récolter un discours spontané semi-dirigé.

Une fois ce corpus d'épreuves orales enregistré, nous avons demandé au sujet de remplir les deux grilles de questionnaires que nous avons choisies. L'EORCT QLQ H&N35 (*European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire Head and Neck Module*) évalue la qualité de vie après un cancer ORL. Le sujet devait cocher la fréquence à laquelle les faits décrits par chaque item lui sont arrivés au cours de la semaine passée, reflétant l'état du patient au moment de l'enregistrement (van Sluis *et al.*, 2018). La cotation était ordinale avec 1 pour Pas du tout, 2 pour Un peu, 3 pour Beaucoup et 4 pour Toujours. Cinq items ont une réponse binaire avec 1 pour Oui et 2 pour Non. Ce test comporte plusieurs dimensions thématiques : la douleur (5 items), la déglutition (5 items), l'altération des sens (2 items), la difficulté de parole (4 items), la mâchoire (5 items), l'alimentation en public (3 items), la toux (1 item), le contact social (4 items), l'altération de la sexualité (2 items) et le mal-être (2 items). Le VHI (*Voice Handicap Index*) permet de connaître le handicap ressenti par les patients de l'étude et le mesure à trois niveaux : fonctionnel, physiologique et émotionnel (Jacobson *et al.*, 1997). Les scores obtenus permettent d'attribuer des grades (*mild, moderate, severe*) à chaque catégorie du test selon des écarts-types étalonnés (Jacobson *et al.*, 1997).

Pour structurer notre méthode d'analyse, nous avons sélectionné des paramètres rythmiques globaux d'une part, et locaux d'autre part, ces derniers constituant l'un des aspects novateurs de notre étude. Les paramètres rythmiques tels que décrits par Canault (2017) que nous avons définis comme globaux sont le TMP, le Temps de parole (durée totale d'un l'énoncé), la Vitesse de parole et le Débit articulatoire. La vitesse de parole correspond au nombre de syllabes prononcé par seconde et se calcule à partir du Temps de parole. Le débit articulatoire se calcule à partir du Temps total de locution, c'est-à-dire le temps de parole auquel on enlève le temps de pause. Comme la situation de communication est un facteur qui a une grande influence sur la vitesse de la parole, notre corpus comporte trois contextes d'évaluation de ces paramètres : la lecture de texte, le discours spontané semi-dirigé et les diadococinésies. Pour ces dernières, notre propos étant de confronter la vitesse de parole aux vitesses mesurées dans les autres situations (lecture et spontané) ; nous n'avons donc pas converti ce débit en nombre de phonèmes par seconde, contrairement à ce qui est prévu par le protocole Monpage (Pernon *et al.*, 2020).

Nous avons ensuite étudié des paramètres rythmiques que nous avons définis comme locaux, c'est-à-dire qui s'intéressent non pas à un énoncé dans son ensemble, mais à des segments de parole. Ils regroupent le nombre de rhèmes, le nombre de syllabes par

rhèse, le nombre total de pauses, le nombre de pauses silencieuses, le nombre de pauses sonores ainsi que des mesures de durée et de variabilité des voyelles. Deux composantes du corpus ont été dédiées à ces mesures. La lecture de la phrase « Mélanie vend du lilas » avec et sans intonations. Nous voulions rechercher une éventuelle modification des paramètres mesurés pour la même phrase, lue sans intonation particulière. Enfin, nous avons analysé ces paramètres pour l'ensemble des phrases ambiguës. Comme ces phrases nécessitent, pour être distinguées, de respecter un nombre de rhèses et de pauses imposé par la syntaxe, elles nous ont permis d'identifier la capacité ou non d'un sujet à respecter les contraintes rythmiques syntaxiques en lecture. Pour cela, il nous a fallu calculer la différence entre le nombre de pauses de la première phrase du couple et le nombre de pauses de la deuxième. Notons que la vitesse d'enchaînement des phonèmes participe également à la distinction des phrases mais dans le cadre de notre étude, nous nous sommes concentrés sur les pauses. Nous considérons la différence (de nombre de pauses entre chaque phrase du couple, mesurée en moyenne) chez les sujets sains comme conventionnelle et nous l'avons comparée aux différences (moyennes) obtenues pour les patients (en distinguant VO et VTO). Cette méthode est appliquée pour chacun des six couples de phrases.

Comme préconisé dans une étude portant sur l'analyse de la parole après LT (Most, Tobin et Mimran, 2000), nous nous sommes intéressés aux durées brutes et relatives de durée des voyelles d'un énoncé. Pour cela, nous avons repris les méthodes utilisées par Mairano et Romano (2011) et Pillot-Loiseau et Xie (2018) : il s'agit de calculer le pourcentage de temps de prononciation de l'ensemble des voyelles d'un énoncé, $V_{perc} = t_v / (t_c + t_v)$ noté $V\%$, ainsi que la variabilité du temps de prononciation des voyelles au sein de la phrase, $V_{dev} = \sigma_{tv}$ (où t_v et t_c sont les temps respectifs de prononciation des voyelles et des consonnes, et σ_{tv} est l'écart-type du temps de prononciation des voyelles au sein de l'énoncé) noté ΔV . Le traitement de ces données a nécessité l'élaboration de fichiers représentant une segmentation de l'énoncé par phonème. Ainsi, à partir de fichiers .txt comprenant l'orthographe du contenu du son, et du son correspondant, le site WebMaus (Kisler, Reichel et Schiel, 2017) a permis de générer des fichiers .TextGrid lisibles sous Praat avec la segmentation phonème par phonème de notre corpus. Comme WebMaus n'est pas programmé pour les voix pathologiques, nous avons donc corrigé la segmentation manuellement via le logiciel Praat (exemple en Annexe B) pour la lecture du couple de phrases ambiguës n°5 par un sujet VO). Ces fichiers sont ensuite insérés dans le logiciel Correlatore (Mairano et Romano, 2011) pour extraire les données $V\%$ et ΔV .

Nous avons formulé trois hypothèses qui guideront notre étude. Elles s'accompagnent d'une méthodologie statistique qui servira en partie à y répondre. L'hypothèse (1) suppose une influence de la LT sur les paramètres rythmiques globaux de la parole, comme la vitesse articulatoire en spontané, la lecture et les diadococinésies, avec une différence plus marquée pour la diade VO-T que pour VTO-T. L'hypothèse (2) suppose l'observation d'un effet de la LT sur les paramètres rythmiques locaux de la parole tels le nombre de rhèses, le nombre de pauses et les marqueurs ΔV et $I\%$, avec une différence plus marquée pour la diade VO-T que pour la VTO-T. En hypothèse (3), nous avons supposé une éventuelle corrélation du VHI avec l'EORTC QLQ H&N 35 et les paramètres rythmiques globaux (qui représentent l'état général d'un énoncé) : nous avons voulu rechercher un lien entre l'état de la parole objectivé par des mesures et l'état vocal ressenti. Pour les deux premières hypothèses, nous avons donc comparé deux à deux les sujets VO et T ainsi que les sujets VTO et T. Pour les hypothèses (1) et (2), comme l'échantillon était inférieur à trente sujets par groupe, nous avons choisi un test statistique non-paramétrique. Nous avons comparé les valeurs obtenues entre deux groupes, deux fois : VO vs. T et VTO vs. T, et ce avec des échantillons indépendants. Nous avons donc utilisé un test de *Mann & Whitney*. Pour notre hypothèse (3), nous avons également choisi un test non-paramétrique. Comme nous voulions chercher ici une corrélation statistique entre 2 variables pour un seul groupe (VO et VTO réunis), nous avons décidé d'utiliser un test de corrélation de *Spearman*. Pour l'ensemble de nos hypothèses, le critère de validité appliqué est une p-value inférieure à 0,005 (test de *Fisher*). D'après Laccourreye *et al.* (2019), le test de Fischer permet de considérer nos résultats comme « significatifs » (p-values inférieures à $p = 0,005$), « suggestifs » ($0,005 < p < 0,05$) ou « non-significatifs » ($p > 0,5$).

Résultats

Notre population effectivement étudiée a regroupé pour le groupe VO de 9 hommes et 3 femmes, pour le groupe VTO de 15 hommes et 2 femmes, et pour le groupe T de 4 hommes et 8 femmes. Nos deux groupes de patients VO et VTO présentaient une répartition assez similaire en âge et en sexe. Nous notons que la proportion de femmes du groupe T était nettement supérieure aux deux autres (67% vs. 25% et 12%) et comportait une plus grande hétérogénéité de l'âge avec des sujets de 42 à 87 ans vs. 55 à 84 pour la VO et 57 à 83 pour la VTO. Un test de *Mann & Whitney*, réalisé pour évaluer la différence d'âge et de sexe entre les groupes, a donné les résultats suivants : une différence de sexe significative entre les groupes VTO et T ($Z = -2,98439$, $p = 0,0029$), suggestive entre les

groupes VO et T ($Z = -1,97$, $p = 0,048$) et non-significative entre les deux groupes de patients ($Z = -0,87$, $p = 0,379$). Les différences d'âges entre les groupes sont toutes non-significatives ($Z = -1,50$ et $p = 0,132$ pour VO vs. T, $Z = -1,22$ et $p = 0,222$ pour VTO vs. T, et $Z = 0,63$ et $p = 0,525$ pour VTO vs VO).

Nous avons mesuré le temps de parole (en secondes) de la lecture du texte *La bise et le soleil*. Pour le groupe de patients utilisant la VO, les résultats sont très disparates, allant de 48,5 secondes pour le plus rapide à 139 secondes pour le plus lent. Les sujets des deux autres groupes sont moins variables avec des résultats dans un intervalle plus petit : de 41,4 s à 70 s pour le groupe VTO et de 34,88 s à 58,18 s pour le groupe T. La médiane du groupe T (46,76s) est nettement inférieure à celle du groupe VO (73,94s) et inférieure à celle du groupe VTO (52,07s). Cela nous montre que les VTO lisent un peu moins vite que les T mais que les VO lisent beaucoup moins vite que les T (et que les VTO *a fortiori*). Les temps de parole des discours spontanés semi-dirigés varient de 3,06s à 39,07s et ce grand écart rend les calculs de vitesse délicats à interpréter. Les diadococinésies utilisées ici pour le calcul de la vitesse de parole durent chacune 5 secondes, le temps de parole de cette composante a donc été le même pour chacun des groupes.

A partir de tous ces temps de parole chronométrés, nous avons calculé la vitesse de parole qui se mesure en nombre de syllabes par seconde (s/s). Les résultats de la figure 1 (vitesses de parole des patients VO et VTO, comparées au groupe témoin T) montrent une différence entre chaque diade de séries comparées (rejet de l'hypothèse nulle selon laquelle les séries sont identiques). La vitesse est effectivement plus lente pour les deux groupes de patients que pour les témoins en lecture (teintes claires du diagramme), un résultat nettement plus lent pour les patients en moyenne (2,53 s/s pour la VO, 3,40 s/s pour la VTO et 4,08 s/s pour les témoins) et en particulier pour le groupe VO (différence significative avec le groupe T avec ($Z = -3,72$; $p = 0,0002$). La vitesse de parole en spontané (teinte medium du diagramme) est également ralentie pour les patients VO avec une moyenne de 2,69 s/s vs. 3,39 s/s pour le groupe T même si la différence entre ces deux groupes n'est pas significative (différence non-significative entre le groupe VO et le groupe T ($Z = -1,58$; $p = 0,112$). En revanche, la moyenne de la vitesse de parole en spontané est légèrement plus élevée pour le groupe VTO avec une moyenne de 3,84 s/s. Le discours spontané a donné des résultats très différents. Un test de *Mann Whitney* a été réalisé pour comparer les différentes vitesses mesurées. La comparaison entre les groupes VO et T donne une valeur de p significative en lecture ($Z = -3,72$; $p = 0,0002$) et en diadococinésies ($Z = -3,58$; $p = 0,0003$) et une valeur de p non-significative en spontané ($Z = -1,58$; $p = 0,112$). Le même test est réalisé entre les groupes VTO et T avec une valeur

de p suggestive pour la lecture ($Z = -2,76$; $p = 0,005$) et les diadococinésies ($Z = 2,27$; $p = 0,014$) et une valeur de p non-significative en spontané ($Z = 0,04$; $p = 0,296$). Nous présentons dans la figure 1, les valeurs des débits de parole des trois groupes (T en gris, VTO en bleu et VO en rouge).

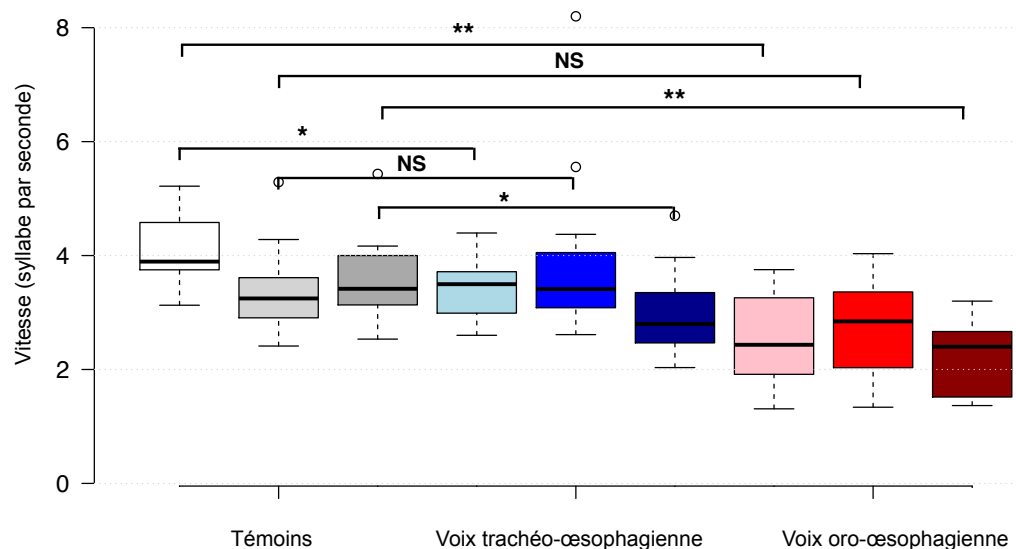


Figure 1. Distribution de la vitesse de parole en fonction du groupe. Les teintes de couleur correspondent à une situation d'analyse différente (claire : lecture, medium : discours spontané, foncée : diadococinésies). Les barres indiquent le niveau de significativité – avec NS pour non-significatif (soit $p > 0,05$), * pour suggestif (soit $0,05 > p > 0,005$) et ** pour significatif (soit $p < 0,005$) – pour chaque comparaison.

Nous avons relevé par ailleurs une différence de vitesse de parole spontanée entre les deux groupes de patients, dont la significativité est suggestive ($Z = 2,39$; $p = 0,016$). Cette comparaison (VO vs. VTO) ne faisait pas partie de nos hypothèses mais peut être une information pertinente : VTO et VO étant très proches, l'impact de la LT sur la vitesse de parole en spontané (uniquement) ne se manifesterait que pour les patients utilisant la VO. La vitesse de parole en diadococinésies est plus lente pour les patients que pour le groupe T. D'une part, la différence du nombre d'occurrences (d'un item en 5 secondes) est fortement marquée et significative entre VO et T ($Z = -3,58$; $p = 0,0003$) avec des moyennes allant de 4,04 occurrences (item 4 « cla cla cla ») à 5,17 (item 6 « cla tra ») pour la VO, et de 7,13 (item 5 « tra tra tra ») à 7,83 (items 1 « ba ba ba » et 6 « cla tra ») pour le groupe T. La différence existe aussi entre VTO et T mais elle est moins importante (moyennes de 6,06 (item 5 « tra tra tra ») à 7,06 s/s (item 6 « cla tra ») pour la VTO) et avec un niveau de significativité suggestif ($Z = 2,27$; $p = 0,014$).

L'analyse de la vitesse de parole à partir des diadococinésies nous a également permis d'identifier que pour l'ensemble des groupes (VO, VTO et T), les items « ba ba », « cla tra » et « de de de » sont les trois premiers qui sont prononcés le plus

rapidement. Ensuite, le groupe VO a eu, dans l'ordre décroissant de vitesse, les items : « go go go », « tra tra tra » et « cla cla cla » ; le groupe VTO : « cla cla cla », « go go go », « tra tra tra » et le groupe T : « go go go », « cla cla cla », tra tra tra ». On peut observer que l'item « cla tra » est dans la moyenne haute de vitesse pour tous les groupes (5,17 s/s pour la VO, 7,06 pour la VTO et 7,83 pour le groupe T). Or, les autres items de pattern syllabique CCV se trouvent plutôt dans la moyenne basse. Nous pouvons donc penser que c'est davantage le fait que l'item « cla tra » ne comporte que deux syllabes qui nous a permis de comptabiliser plus d'occurrences que pour le pattern syllabique CCV. Sur les items à trois syllabes (les 5 autres), on remarque que les patterns syllabiques CV sont prononcés plus rapidement et les patterns syllabiques CCV plus lentement (pour tous les groupes).

Nous avons manuellement compté le nombre de rhèses et le nombre de pauses nécessaires à la lecture de la phrase « Mélanie vend du lilas » ainsi que le nombre moyen de syllabes par rhèse (dépendant du nombre de syllabes total de la phrase). Pour cette partie du corpus (lecture de phrase simple), aucun sujet VO n'a produit de pauses silencieuses et sonores mais seulement l'une ou l'autre. Les résultats très élevés pour le groupe VO (de 0 à 6) et très bas (0) pour les groupes VTO et T révèlent une différence notable avec d'un côté le groupe VO et de l'autre les groupes VTO et T. Il en va de même pour les rhèses (1 à 7 pour VO et 1 pour VTO et T). La lecture des phrases ambiguës nous a permis de relever la différence entre le nombre de pauses total effectuées entre chaque couple de phrase, puis de comparer cette différence pour chaque groupe (résultats en Annexe A) : nous avons en effet comparé la différence du nombre de pauses entre les deux phrases, pour chaque diade (VO/T, VTO/T). Nous avons opéré une soustraction entre le nombre de pauses de la première phrase du couple, et le nombre de pauses de la deuxième phrase du couple pour savoir comment le sujet a modifié son énoncé (d'un point de vue des pauses) pour distinguer les deux phrases. Ensuite, nous avons gardé cet indice de différence pour comparer sa valeur entre les différents groupes de patients. La valeur n'est utile qu'en la comparant entre les sujets pour un même couple de phrases. Les résultats pour les couples de phrases n°1 (« Elle manque à mon amie », « Elle manqua mon amie ») et n°3 (« il aime à en mourir », « il aima en mourir ») montrent une différence entre les patients et les témoins (différence pour VO de -0,58 pauses, pour VTO de -0,29 pauses et pour T de -0,33 pauses pour le couple 1 et différence pour VO de -0,08 pauses, pour VTO de -0,31 et pour T de -0,33). Le couple n° 3 montre une différence nettement plus importante pour la VO avec une différence de -0,08 isolée vs. les différences de VTO et T qui sont proches (-0,31 et -0,33). Les autres couples de phrases (n°2, 4, 5, 6) nous ont

donné des résultats relativement aléatoires. Nous avons proposé une démonstration d'analyse de ces phrases (Figure 5 en Annexe A).

Après avoir analysé la problématique des pauses, nous avons calculé le débit articulatoire de la lecture de la phrase « Jean mène un journal ». Ce rapport en syllabe/seconde évalue la rapidité de la parole à partir du temps de locution (sans tenir compte des pauses). Nous avons ensuite comparé les valeurs de ce débit entre les groupes avec un test de *Mann & Whitney*. Nous avons obtenu pour la comparaison VO/T, une différence suggestive ($Z = 2,03$; $p = 0,021$) révélant un débit plus rapide pour le groupe T que le groupe VO ; et pour la comparaison VTO/T, une différence non-significative ($Z = 0,69$; $p = 0,486$) révélant une certaine similarité de débit pour ceux deux groupes. Notons que la comparaison VTO/VO était suggestive ($Z = -2,31$; $p = 0,020$) avec un débit plus rapide pour les VTO que les VO.

La Figure 2 représente les résultats ΔV et $V\%$ obtenus à la lecture des phrases ambiguës n°5. Nous avons choisi ce couple car c'est celui pour lequel les paramètres rythmiques locaux mesurés manuellement étaient le plus similaires entre les groupes. En effet, les médianes du nombre de syllabes par rhème de la phrase sont proches (2,08 pour la VO, 2,50 pour la VTO et 1,66 pour le groupe T), et sont égales pour la phrase 5 bis (3,75 pour la VO et 5 pour VTO et T). Ainsi, c'est le couple de phrase pour lequel la plupart des sujets VO a pu réaliser le même enchaînement de rhèmes et de pauses que les autres groupes (exemple d'une segmentation des phrases 5 et 5bis en Annexe B). Le graphique de gauche se rapporte à la phrase ambiguë n° 5 « Jean mène un journal » et celui de droite à la phrase ambiguë n°5 bis « J'emmène un journal ». Nous avons représenté l'évolution de ΔV en fonction de $V\%$ ainsi que l'écart-type de ces valeurs.

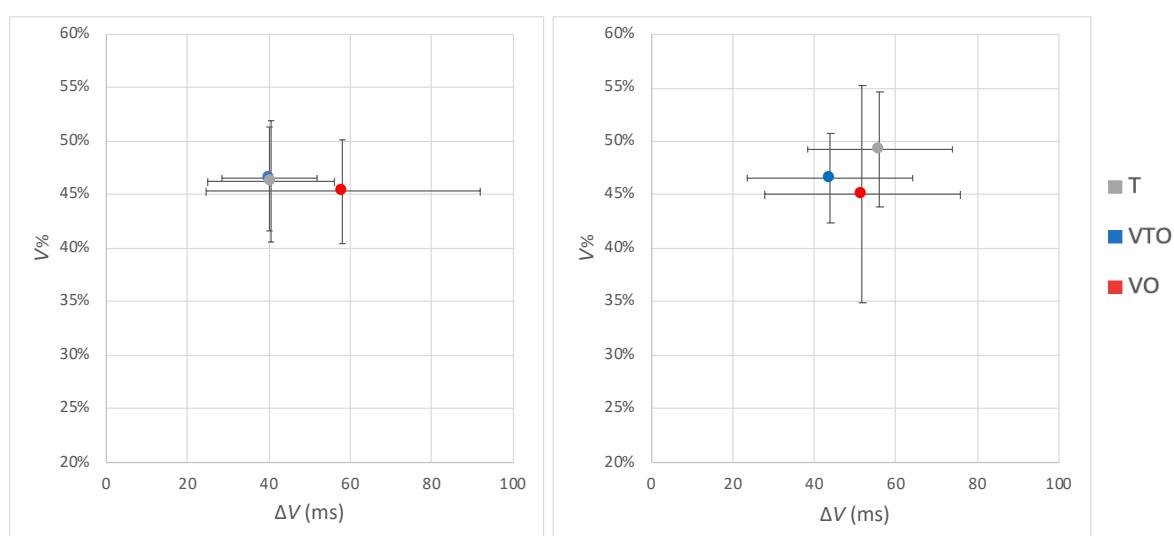


Figure 2. Durée d'émission de voyelle en pourcentage ($V\%$) en fonction du quotient de variabilité de la durée des voyelles (ΔV), avec les barres d'erreur représentant les écarts-types. Gauche : *Jean mène un journal*. Droite : *J'emmène un journal*.

Pour « Jean mène un journal » (gauche), les VO se distinguent des VTO et T par une plus grande variabilité de durée des voyelles (variabilité de durée étant non-significativement plus importante que pour les VTO et témoin). Pour « J’emmène un journal », les trois groupes sont distincts avec plus de durée et plus de variabilité de durée pour les T (non-significativement plus importante que pour les groupes VO et VTO). Enfin, on remarque que les VO et les VTO ont moins changé leurs valeurs entre les deux phrases que les T. On a donc par l’observation de cette figure, une plus grande modification des paramètres ΔV et $V\%$ entre les deux phrases pour le groupe T que pour les deux groupes de patients, différence non-significative. Par ailleurs, on remarque que les barres d’erreurs sont plus grandes pour la VO, indiquant une plus grande variabilité inter-sujet.

Au vu des résultats analysés précédemment qui démontrent certaines difficultés de parole pour les patients, il convient à présent d’évaluer le lien entre ceux-ci et nos échelles de qualité de vie. Les questionnaires de qualité de vie et de handicap vocal ont été distribués aux deux groupes de patients uniquement. Nous avons obtenu au VHI les résultats présentés sur la Figure 3. Les scores bruts sont donnés en Annexe C.

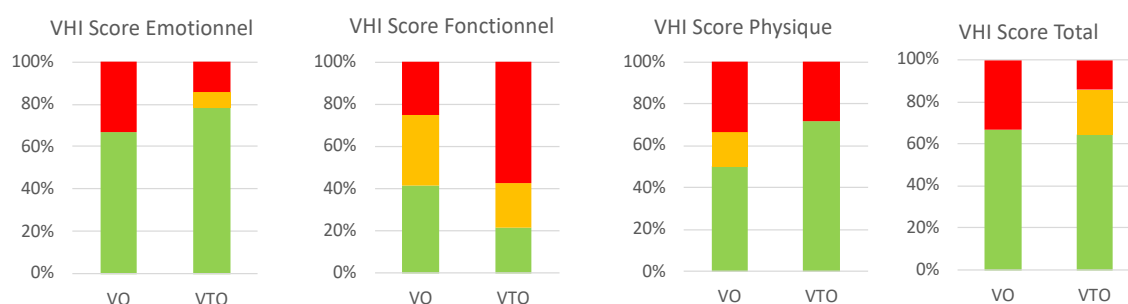


Figure 3. Répartition des scores de chaque sous-partie du VHI selon le Groupe (VO ou VTO) dans différents degrés de sévérité (calculés à partir de l’étalonnage réalisé par Jacobson *et al.*, 1997) avec stade léger (Mild) en vert, stage modéré (Moderate) en jaune et stade sévère (Severe) en rouge.

Nous avons pu constater que les degrés de sévérité des VO sont relativement proches pour le score émotionnel et le score total du VHI ainsi que pour les VTO à l’exception du nombre de stades modérés. On peut remarquer que le groupe VO comporte systématiquement plus de stades sévères, excepté pour le handicap vocal fonctionnel qui est ressenti plus sévèrement par les sujets VTO. On constate l’absence du degré modéré dans trois colonnes (score émotionnel VO, score émotionnel VTO et score total VO). Pour résumer les résultats à l’EORTC QLQ H&N35, nous avons calculé les moyennes des scores de chaque dimension, et ce pour chaque groupe de patients. Nous avons donc obtenu des scores moyens (les réponses possibles allant graduellement de 1 à 4 comme expliqué dans la partie méthode). Les moyennes des scores bruts sont représentées en Annexe C. Nous avons réalisé un tableau récapitulatif des résultats du test de *Spearman*

(Tableau 1), mettant en évidence une corrélation des résultats aux questionnaires qualitatifs entre eux.

Tableau 1 : Tableau des corrélations, évaluées avec le rho de Spearman, entre les dimensions de l'EORTC QLQ H&N 35 et le VHI. Les p-values en gras représentent les corrélations significatives.

Variables	ρ de Spearman	p-value
VHI Total / Sens	0,5124	0,124
VHI Total / Alimentation public	0,5510	0,0079
VHI Total / Contact social	0,5620	0,0053
VHI Total / Parole	0,6505	0,0008
VHI Fonctionnel / Sens	0,6504	<0,0001
VHI Fonctionnel / Toux	0,4226	0,0128
VHI Fonctionnel / Parole	0,3728	0,0356
VHI Émotionnel / Alimentation public	0,6191	0,0022
VHI Émotionnel / Sens	0,4300	0,0406
VHI Émotionnel / Contact social	0,6335	0,0012
VHI Émotionnel / Parole	0,5225	0,0105
VHI Physique / Toux	0,3820	0,0309
VHI Physique / Sens	0,6189	0,0002
VHI Physique / Parole	0,3728	0,0360

Le tableau 1 montre un lien entre la qualité de vie après un cancer ORL et le handicap vocal ressenti. Certaines corrélations comme le VHI émotionnel/Contact social, VHI physique/ Sens, VHI total / Parole, semblent évidentes mais méritent d'être objectivées. La Figure 4 montre les scores totaux du VHI en fonction de la vitesse de parole en lecture et en spontané.

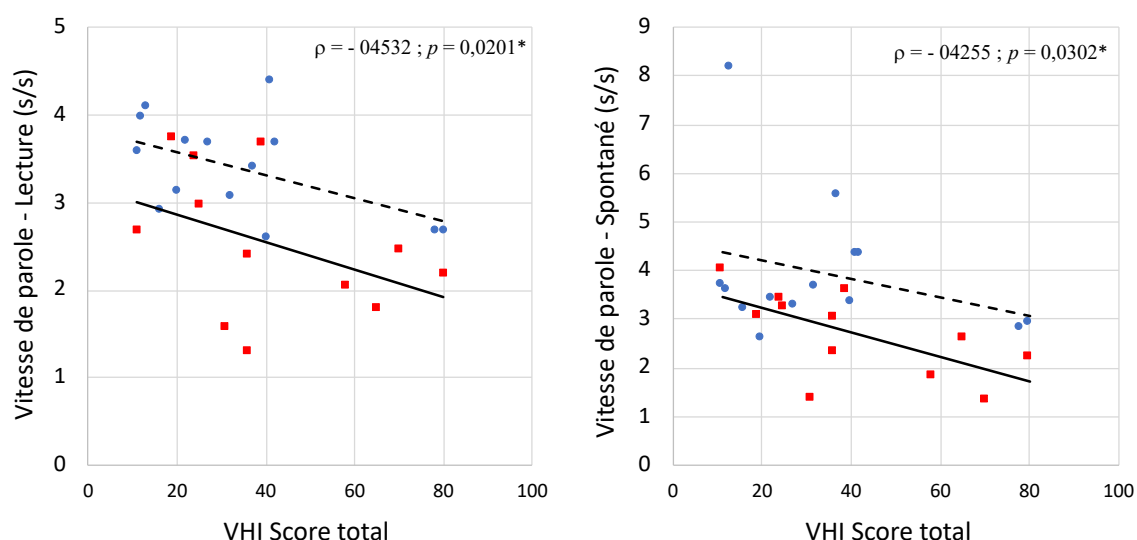


Figure 4. Nuage de points et courbe de régression linéaire. Gauche : vitesse de parole en lecture en ordonnées, score total au VHI en abscisses. Droite : vitesse de parole en spontané et score total au VHI en abscisses. Pour les deux figures sont représentés les groupes VO (points bleus) et VTO (carrés rouges) ; les lignes pleines représentent la courbe linéaire du groupe VO et la ligne en pointillé la courbe linéaire VTO.

Cette figure illustre les corrélations testées. Nous avons découvert une corrélation négative de significativité suggestive entre la vitesse de parole en lecture et le VHI total ($\rho = -0,45$; $p = 0,0201$) ; et une corrélation négative de significativité suggestive vitesse de parole en spontané / VHI ($\rho = -0,425$; $p = 0,0302$). Ces différences ne sont donc pas significatives mais nous ont permis de mettre en évidence une tendance (représentée par les courbes linéaires sur la figure) pour chacun des groupes de patients, plus les vitesses de paroles sont élevées, plus le score de VHI est bas. La corrélation est plus marquée entre le VHI et vitesse de parole en lecture (figure 4 à gauche) que celle entre le VHI et vitesse de parole en spontané (figure 4 à droite) : différence soutenue par les deux valeurs de p (la corrélation étant plus importante pour la lecture que pour le spontané). Ceci peut s'expliquer par le fait que les mesures de débit spontané, comme nous l'avons vu, ont donné des résultats très variables (Figure 1).

Discussion

Sur l'ensemble de nos résultats, nous avons pu constater que les paramètres locaux et globaux de la parole étaient effectivement impactés par la LT. Certaines différences VO/T et VTO/T que nous avons résumées ici sont significatives malgré le faible nombre de sujets testés. Nous avons également pu identifier des paramètres rythmiques en lien avec les scores au VHI sur l'ensemble des patients ainsi que des corrélations entre des items du VHI et de l'EORTC QLQ H&N 35. Nous avons observé une tendance pour les groupes VO et VTO à avoir une parole plus lente et ce dans les paramètres globaux comme locaux. Notre étude n'a révélé qu'une légère altération des paramètres temporels évalués pour le groupe VTO. En effet, les résultats obtenus pour ce groupe sont soit identiques au groupe témoin, soit légèrement différents. En revanche, comme nous le supposions, il semblerait que la VO présente des paramètres temporels notablement impactés. Les résultats obtenus pour ce groupe ont été systématiquement différents du groupe témoin avec pour la plupart des résultats, une grande différence. Selon nos analyses, l'utilisation de la VTO permettrait de produire une parole proche de la parole laryngée (d'un point de vue temporel) et l'utilisation de la VO ne le permettrait pas. Nous retenons en particulier que (1) la vitesse de parole en tâche de lecture et en diadococinésies a donné des résultats mettant en évidence des différences significatives entre les groupes VO et T mais pas pour la tâche de discours spontané. Pour la comparaison de ces mêmes paramètres, aucune différence significative n'a été identifiée entre le groupe VTO et le groupe T. (2) Le calcul des pauses et des rhèmes a révélé une différence flagrante entre les groupes (avec d'un côté VTO et T, et de l'autre VO) mais trop imprécise pour être prise en compte d'un point de

vue statistique. (3) Les résultats obtenus pour les paramètres locaux relatifs à la voyelle (ΔV et $V\%$), s'ils n'ont pas mis en évidence de différence significative, ont tout de même permis d'identifier des caractéristiques telles que : de plus grands écarts-types (reflétés par les barres d'erreur sur notre figure) pour les sujets VO et un chevauchement des valeurs entre VTO et VO pour une des deux phrases. (4) Nous avons trouvé des corrélations entre certains aspects évalués de la qualité de vie et du handicap vocal ainsi qu'une tendance de l'ensemble des patients de l'étude à avoir un handicap vocal ressenti relatif à sa vitesse de parole (en lecture et plus modérément en spontané).

A l'instar de Grosman (2018), nous avons généré un corpus composé de plusieurs « situations communicatives » différentes. Nous avons différencié nos trois situations (lecture, spontané et diadococinésies) pour la vitesse de parole et obtenu des résultats intéressants mais nous n'avons pas pu comparer le nombre de pauses entre les situations car notre épreuve de discours spontané n'était pas suffisamment dirigée (temps de parole trop différents). Le sujet devait nous raconter ce qu'il faisait avant de venir (souvent un déjeuner ou un trajet), épreuve dont la longueur est très dépendante de la personnalité. Il en a donc résulté des temps de parole allant de 3,16 à 39,07 secondes. Pour s'adapter à cette hétérogénéité des temps de discours spontanés, nous avons donc décidé de garder uniquement la vitesse de parole (car elle est représentée par un rapport), avec une grande réserve d'interprétation. Pour autant, nous avons tout de même pu décrire des caractéristiques pausales en exploitant la situation de lecture. Contrairement à l'analyse de Grosman (2018) qui détermine une pause silencieuse en tant que telle à partir de 250 ms, nous n'avons pas associé une pause à un temps. Les pauses silencieuses de notre étude ont été sélectionnées si elles étaient premièrement identifiées comme telles à l'oreille de l'examineur, et deuxièmement objectivées par un silence visible sur le signal acoustique. De plus, nous n'avons pas repris sa catégorisation des pauses qui stipule globalement que les pauses courtes sont dues à des éléments articulatoires, et les pauses longues à des éléments segmentaux discursifs. En effet, ces catégorisations sont formulées à partir d'un corpus comportant des épreuves élaborées (discours, politique, homélie etc.) qui ne sont pas assimilables à celles de notre corpus, et difficilement applicables à des paroles pathologiques comme celle de notre étude. Les mesures de pauses en lecture simple que nous avons effectuées, donnent des résultats très élevés pour le groupe VO (de 0 à 6) et minimaux (0) pour les groupes VTO et T. Elles démontrent que ce groupe de patients aura davantage de difficultés à se faire comprendre que les VTO, car, d'après l'étude de Motta (2001) (spécifique à la LT) - qui identifie chez les patients VO la réduction d'intelligibilité comme la conséquence d'une fluence de parole inadéquate - plus le nombre de pauses

inappropriées est réduit (voire absent), meilleure est l'intelligibilité. De plus, la grande différence du nombre de pauses et de rhèmes entre le groupe VO et les groupes VTO et T va dans le sens de notre hypothèse (2) - qui prévoit un nombre plus élevé de pauses pour la VO. Cependant, des résultats très différents entre les groupes avec des égalités au sein des groupes VTO et VO, mettent en évidence un manque de sensibilité de l'épreuve de lecture simple de notre corpus. Ces mêmes mesures pausales nous ont permis d'identifier, dans la lecture de phrases ambiguës, la capacité des sujets à respecter les contraintes syntaxiques de pauses - nécessaires pour distinguer deux phrases homophones. Selon la modélisation des pauses de Grosman (2018), il y a une forte prédiction syntaxique de la pause. En effet, nous avons retrouvé une similarité fréquente du nombre de pauses (pour la même phrase lue) au sein des groupes, et entre les groupes, sauf pour la VO. Ces phrases nous ont permis de supposer un nombre « moyen » de pauses syntaxiques différent entre les deux phrases du couple, participant à les distinguer. Il s'agirait donc, pour la problématique rythmique des patients LT qui nous intéresse, de comparer la différence du nombre de pauses entre les deux phrases, pour chaque diade (VO/T, VTO/T). Nous avons obtenu des résultats relativement différents sur 2 couples de phrases mais des résultats relativement aléatoires sur d'autres. Cela peut être dû au choix de cette épreuve qui comporte une grande contrainte temporelle syntaxique, facteur extrêmement cadrant et permettant peut-être aux sujets pathologiques de s'exprimer comme les sujets contrôles, donnant des résultats très proches. Comme l'explique Van Rossum (2008), on distingue chez les sujets VO les pauses d'injection des pauses prosodiques dite « appropriées » (celles voulues par la syntaxe). Il explique aussi que la fréquence des pauses d'injection (inappropriées) augmente avec la longueur de l'énoncé lu et la longueur des mots de l'énoncé lu. Nous avons pu, grâce à l'analyse des phrases ambiguës (vue précédemment), savoir si le nombre de pauses inappropriées augmentait avec la longueur de l'énoncé. Ainsi, nous avons pu voir si la différence entre les groupes de patients et le groupe T varie avec la longueur de l'énoncé. Nos mesures de nombre de pauses n'ont cependant pas permis d'identifier un rapprochement entre la fréquence des pauses inappropriées et la longueur de l'énoncé. Une comparaison intéressante nous a permis de déterminer que les pauses ne sont pas la seule cause de ralentissement du débit de la parole après LT. Effectivement, les résultats de débit articulatoires (débit sans le temps de pause) montrent une différence suggestive entre les VO et les deux autres groupes. Les différences obtenues pour la vitesse de parole (débit avec le temps de pauses) sont significatives donc néanmoins plus importantes que pour le débit articulatoire. Nous pouvons donc conclure que d'après nos résultats, les pauses seraient un facteur majeur de ralentissement de la rapidité de l'élocution des patients

utilisant la VO mais que même lorsque les pauses n'entrent pas en compte, leur parole est nettement (suggestivement) plus lente que les sujets contrôles (et que les sujets VTO en l'occurrence).

Nos résultats obtenus pour les mesures autour de la voyelle (ΔV et $V\%$), ne participent que partiellement à corroborer notre deuxième hypothèse de recherche. Nous pouvons émettre la possibilité que l'absence de différence significative de la durée des voyelles entre les patients VO et les témoins est due à une faiblesse du nombre de sujets évalués. En effet, Most *et al.* (2000) permettent d'identifier par un test *Anova* (rendu possible par un plus large échantillon de sujets) une différence significative de ces variables entre la VO et la voix laryngée.

Notre recherche autour des échelles qualitatives est innovante car elle permet de relier par des données chiffrées, des ressentis subjectifs évalués. Une des corrélations trouvées a révélé que le retentissement de difficultés de parole sur la qualité de vie (mesuré par l'EORTC QLQ H&N 35) était lié au handicap vocal ressenti (évalué grâce au score total du VHI). Notre étude quantitative portant elle-même sur la parole, cette corrélation vient établir un lien inattendu et rarement évalué dans les études d'analyses temporelles de la parole des LT de la littérature. De plus, les tendances que nous avons observées (Figure 4) montrent un lien encore plus spécifique : ils établissent une corrélation entre nos résultats qualitatifs et nos mesures rythmiques. Par ailleurs, nous avons découvert des corrélations entre des dimensions de qualité de vie et de handicap vocal ressenti. Entre autres, nous avons appris que les patients des deux groupes réunis présentaient (1) un trouble fréquent des sens évalués (odorat et goût) correspondant à un handicap vocal de type fonctionnel ressenti et (2) une difficulté à s'alimenter et à avoir des contacts sociaux reliée à un handicap vocal de type émotionnel ressenti. Ces résultats ne sont pas directement en rapport avec le sujet de notre étude mais sont relativement sérieux et apportent des informations utiles pour prendre en compte globalement les effets de la LT sur les patients que nous étudions. Une information supplémentaire découverte lors de l'analyse des résultats du VHI tempère le postulat que la VO est toujours plus « impactée » que la VTO. Effectivement, le handicap vocal fonctionnel est ressenti beaucoup plus sévèrement par les VTO que les VO, ce qui peut s'expliquer par le fait que la fonctionnalité de la VTO est dépendante notamment de la prothèse phonatoire, qui peut faire défaut.

Nos conclusions sont tempérées par un certain nombre de biais que nous connaissons. Premièrement, au niveau de la population. Le recrutement des patients a été une étape difficile de notre étude qui a demandé de nombreuses prises de contact, souvent

infructueuses, et de déplacements. Par chance, 6 mois ont suffi à récolter un nombre notable de données qui ont abouti aux résultats que nous avons présentés. Néanmoins, les différences d'âge et de sexe entre nos groupes de patients et le groupe témoin est un facteur de confusion que nous connaissons et qui nous appelle à modérer l'interprétation de nos résultats. Deuxièmement, au niveau de l'organisation de passation du protocole, il aurait fallu proposer aux patients une lecture silencieuse du protocole avant de les enregistrer pour éviter le facteur de confusion d'un mauvais déchiffrement. Ce biais a néanmoins été limité car pour chaque composante du corpus, les sujets devaient réaliser deux lectures à voix haute au minimum de sorte que nous ayons toujours deux échantillons (de la même épreuve) enregistrés. Certains biais n'ont pas pu être écartés par manque d'information récupérée auprès des patients : le type de cancer (protéine P16 ou alcool/tabac), la date d'opération de la LT, savoir si la LT était la première opération ORL ou non, le traitement radiothérapeutique et/ou chimiothérapeutique ou non et le temps qu'a duré la rééducation orthophonique. Ensuite, l'analyse des résultats par la manipulation manuelle des logiciels introduit un biais dans nos mesures qui est réduit au maximum par le traitement des données par une seule personne.

Nous avons pris la décision de dresser un état des lieux des aspects rythmiques de la parole après LT, et donc de réaliser un certain nombre de mesures différentes. Cela nous a certes permis de repérer des signes intéressants dans plusieurs domaines (certains résultats sont même significatifs) mais n'a pas permis de traiter de manière exhaustive un champ de mesures en particulier. Le protocole que nous avons conçu et par conséquent les entretiens, étaient longs (30 minutes, questionnaires compris) et, sur l'ensemble des nombreuses données acquises, une partie n'a pas pu être traitée. Ainsi, les données que nous avons récoltées pourraient servir une nouvelle recherche concentrée sur les pauses en faisant entre autres une caractérisation par critère de durée (brèves, moyennes longues) ou par fonction (respiration, grammaire, hésitation) comme l'a fait Grosman (2018) pour la parole normale. D'autres interprétations des données sont possibles comme l'étude des pauses sonores pour la voix œsophagienne. L'idéal serait de bénéficier d'une analyse perceptive permettant premièrement d'exclure les sujets obtenant moins de 50% d'intelligibilité de leur parole, et deuxièmement de répartir les sujets LT en différents groupes de niveau (bon, médiocre, mauvais), comme l'a fait Motta (2001) ainsi que Most, *et al.* (2000). Une étude perceptive de ce type n'a pas pu être réalisée. Ce biais a néanmoins été limité pour notre étude car nous avons choisi comme critère d'inclusion, un suivi en orthophonie, permettant d'éliminer des sujets ayant une parole non entraînée. Aussi, nous n'avons pas pu acquérir deux mesures très influentes dans le rythme de la

parole (Motta, 2001) pour faute de matériel approprié : les paramètres aérodynamiques comme le volume d'air phonatoire (mesuré en millilitre par des aérophones) ou la pression articulaire (pression intra-orale et débit d'air buccal en contexte de prononciation) (Robieux, Legou, Meynadier et Meunier, 2016). A partir de l'enregistrement de la lecture de phrases ambiguës, il serait pertinent d'explorer les stratégies rythmiques mises en place par les patients LT pour distinguer les deux phrases d'un couple lorsque les pauses ne sont pas un critère majeur de distinction. Enfin, il serait intéressant d'exploiter les données de notre étude pour étudier précisément les phénomènes articulaires des voix après LT (Crevier-Buchman, 2010), notamment pour identifier des comportements propres aux patients LT et découvrir ainsi les compensations articulaires mises en œuvre. Nous savons néanmoins grâce au calcul de la vitesse de parole en situation hors contexte sémantique (diadococinésies), révélant une différence significative entre les scores, que la LT restreint les capacités de vitesse d'articulation « pure » des patients utilisant la VO.

Conclusion

Après LT, les patients doivent développer de nouvelles stratégies d'émission de la parole, alors qu'un système de production optimal et automatique était en place avant l'opération, notamment au niveau du rythme de la parole. Le manque d'intelligibilité et sa répercussion sur la qualité de vie de ces patients justifient un travail de recherche, le but étant de pouvoir donner le plus d'informations possible aux patients et d'apporter des outils d'amélioration de sa voix et de sa parole. Dans ce contexte, notre étude trouve sa place et apporte des éléments conjoints et complémentaires des études déjà réalisées dont nous avons développé quelques exemples.

L'étude que nous avons menée confirme que la LT entraîne une modification du rythme de la parole. Les différents paramètres temporels que nous avons testés ont généré des résultats qui démontrent que les patients LT ont un débit articulaire ralenti, en particulier en lecture. Nous pouvons également affirmer que l'utilisation de la VO entraîne un ralentissement de l'émission de la parole, dans toutes les situations testées (lecture, spontané et diadococinésies). Aussi, nous avons découvert que les patients ressentant un handicap vocal fort avaient également une qualité de vie affectée dans certains domaines comme les sens (goût et odorat) ou le contact social par exemple. Nous avons également pu constater une augmentation du handicap vocal ressenti avec la diminution de la vitesse de lecture. Malgré de nombreux biais que nous avons exposés, l'ensemble de nos analyses soutient nos hypothèses, par l'observation de tendances mais aussi par des résultats statistiques attestés.

Notre étude à petite échelle a d'ores et déjà pu couvrir un nombre important de domaines d'évaluation des critères de durée de la parole. Les informations que nous avons extraites de nos résultats, si elles ne sont pas toutes significatives, peuvent constituer des éléments nouveaux permettant de connaître les tendances de chaque voix de substitution étudiée. Ces éléments pourront être utiles à des patients qui se questionnent sur le choix d'une voix de substitution avant ou après l'opération, mais aussi à des orthophonistes qui chercheraient comment évaluer la vitesse de parole de leurs patients. Aussi, les items de phrases ambiguës en particulier pourront compléter les protocoles de rééducation de perfectionnement de la VO : s'appuyer sur les contraintes syntaxiques pour parvenir à un débit plus « naturel ». Enfin, nos données analysées et non traitées telles que les pauses pourront être exploitées en vue d'un futur travail de recherche.

Bibliographie

Audacity Team (2019). Audacity(R) : Free Audio Editor and Recorder [Computer application]. Version 2.3.3.

Boersma, P. et Weenink, D. (2020). Praat: doing phonetics by computer [Computer program]. Version 6.1.13.

Caelen-Haumont, G. (2006). Prosodie et sens : une approche expérimentale. *Marges Linguistiques*, M.L.M.S., 135-272.

Canault, M. (2017). La phonétique articulatoire du français. Paris : Edition De Boeck Supérieur.

Crevier-Buchman, L., Pfauwadel, M.-C., Chabardes, E., Laccourreye, O., Brasnu, D. et Laccourreye, H. (1991). Etude comparative des paramètres temporels des voix sans larynx (oesophagienne et trachéo-oesophagienne). *Annales d'oto-laryngologie et de chirurgie cervico faciale: bulletin de la Société d'oto-laryngologie des hôpitaux de Paris*, 108, 261–265.

Crevier Buchman, L. (2010). Séquelles articulatoires des laryngectomies partielles., *In Prise en charge orthophonique en cancérologie ORL* dirigé par Giovanni A et Robert D, Solal Edition, Marseille, 43-51.

Crevier-Buchman, L., Vaissière, J. Maeda, S. et Brasnu, D. (2002). Etude de l'intelligibilité des consonnes du français après laryngectomie partielle supra-cricoïdienne. *Rev Laryngol Othol Rhinol* 123(5), 307-310.

Duez, D., Legou, T. et Viallet, F. (2009). Final lengthening in Parkinsonian French Speech Effects of position in phrase on the duration of CV syllables and speech segments. *Clinical Linguistics & Phonetics*, 23(11), 781-793.

Fitch, W. T. (2019). Sequence and hierarchy in vocal rhythms and phonology. *Annals of the New York Academy of Sciences*, 1453(1), 29-46.

Giovanni, A. et Robert, D. (2010). *Prise en charge orthophonique en cancérologie ORL*. Marseille : Solal Editeurs.

Grosman, I., Simon, A. et Degand, L. (2018). Variation de la durée des pauses silencieuses : impact de la syntaxe, du style de parole et des disfluences. *Langages*, 211(3), 13-40.

Jacobson, B. H., Johnson, A., Grywalski, C., Silbergleit, A., Jacobson, G., Benninger, M. S., et Newman, C. W. (1997). The Voice Handicap Index (VHI). *American Journal of Speech-Language Pathology*, 6(3), 66-70.

JMP®, Version <14.1.0>. SAS Institute Inc., Cary, NC, 1989-2019.

Kisler, T., Reichel, U. et Schiel, F. (2017). Multilingual processing of speech via web services. *Computer Speech & Language*, 45, 326–347.

Laccourreye, O, Lisan, Q, Bonfils, P, Garrel, R, Jankowski, R, Karkas, A, Le Boulanger, N, Makeieff, M., Righini, C., Vincent, C. et Martin, C. (2019). Use of P-values and the terms “significant”, “non-significant” and “suggestive” in Abstracts in the European Annals of Otorhinolaryngology, Head & Neck Diseases. *European annals of otorhinolaryngology, head and neck diseases*, 136(6), 469-473.

Le Huche, F., et Allali, A. (2015). La voix sans larynx (5e édition). Marseille : Solal Editeurs.

Mairano, P et Romano, A. (2011). Rhythm metrics for 21 languages. Proc. of XVIIIICPhS (International Congress of Phonetic Sciences), Hong Kong, Chine. 1318-1321.

Motta, S., Galli, I., et Di Rienzo, L. (2001). Aerodynamic findings in esophageal voice. *Archives of Otolaryngology–Head & Neck Surgery*, 127(6), 700-704.

Most, T., Tobin, Y. et Mimran, R. C. (2000). Acoustic and perceptual characteristics of esophageal and tracheoesophageal speech production. *Journal of Communication Disorders*, 33(2), 165–181.

Pernon, M., Levêque, N., Delvaux, V., Assal, F., Borel, S., Fougeron, C., Trouville, R., Laganaro, M. (2020). MonPaGe, un outil de screening francophone informatisé d'évaluation perceptive et acoustique des troubles moteurs de la parole (dysarthries, apraxie de la parole). *Rééducation Orthophonique*, 281, 169-197.

Pillot-Loiseau, C. et Xie, H. (2018). Transfert rythmique du chinois mandarin au français dans l'apprentissage du Français Langue étrangère : Acoustique et perception. *Congrès Mondial de Linguistique Française*, SHS Web of Conferences, 46, 09001.

Robieux C., Legou T., Meynadier Y. et Meunier C (2016) Effort produit et ressenti selon le voisement en français. *XXXI^e Journées d'Étude sur la Parole*, Paris, France. 256-264.

Schindler, A., Mozzanica, F., Ginocchio, D., Invernizzi, A., Peri, A. et Ottaviani, F. (2012). Voice-related quality of life in patients after total and partial laryngectomy. *Auris NasusLarynx*, 39(1), 77-83.

Štajner-Katušić, S., Horga, D., Mušura, M. et Globlek, D. (2006). Voice and speech after laryngectomy. *Clinical Linguistics & Phonetics*, 20(2-3), 195-203.

van Rossum, M.A., Quene, H. et Nooteboom, S.G. (2008). Prosodic boundaries in alaryngeal speech. *Clinical Linguistics & Phonetics*, 22(3), 215-231.

van Sluis, K. E., van der Molen, L., van Son, R. J. J. H., Hilgers, F. J. M., Bhairosing, P. A. et van den Brekel, M. W. M. (2017). Objective and subjective voice outcomes after total laryngectomy: a systematic review. *European Archives of Oto-Rhino-Laryngology*, 275(1), 11–26.

Annexe A

Cette annexe montre des exemples de comparaison de nombre de pauses des phrases ambiguës. Les / indiquent les pauses éventuelles, considérées comme « appropriées » pour chaque phrase.

En gras, on voit la frontière où s'opère le changement syntaxique avec la deuxième phrase.

1 : « Elle **manque à** mon amie. » **1bis** : « Elle **manqua /** mon amie.

2 : « Cet homme /est **énormément** bête. » **2bis** : « Cet homme est **énorme / et m'embête.** »

3 : « Il **aime / à** en mourir. » **3bis** : « Il **aima / en** mourir. »

4 : « Je suis inquiet de **la varice** de ma mère. » **4bis** : « Je suis inquiet / de **l'avarice** de ma mère. »

5 : « **Jean / mène** un journal. » **5bis** : « **J'emmène** un journal. »

6 : « Le **rapace /** la noie. » **6bis** : « Le **rat / passe** la noix. »

Nous avons opéré le calcul suivant (exemple avec le couple n°1) :

Enoncé d'un sujet VO avec 3 pauses non « appropriées » comme suit « Elle manque / à / mon / amie. » Son nombre de pauses pour la phrase n°1 = 3. « Elle / manqua / mon amie. »

Son nombre de pause pour la phrase 1bis = 2. On fait la différence entre les deux nombres de pauses, ce qui nous donne $3 - 2 = 1$. Puis, on fait la moyenne de ces différences pour chaque groupe de patients et pour chaque couple de phrases, ce qui nous donne les résultats présentés dans le graphique ci-dessous.

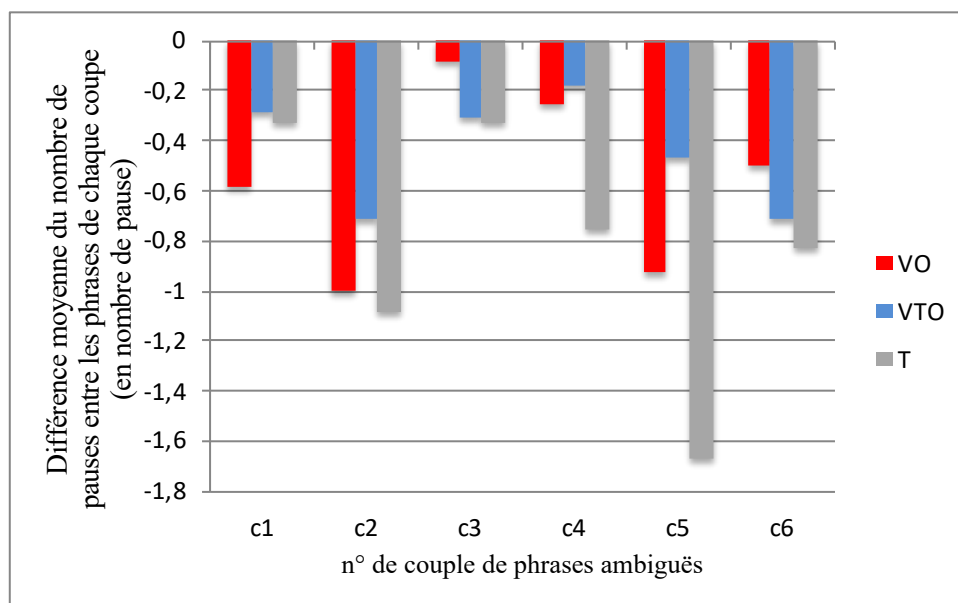


Figure 5. Différence du nombre de pauses entre les phrases de chaque couple de phrases ambiguës (phrase N moins phrase Nbis). Groupe VO (rouge), groupe VTO (bleu) et groupe T (gris). L'axe des ordonnées représente la différence du nombre de pauses moyen qui différencie les phrases de chaque couple et l'axe des abscisses représente les 6 couples de phrases étudiés.

Prenons le cas du couple 1 (haut gauche du graphique) pour lequel on observe une plus grande différence de pauses entre les deux phrases pour la VO (rouge) et une différence plus petite et quasiment égale pour les groupes VTO (bleu) et T (gris). Il est donc possible que les sujets aient eu besoin par exemple de marquer une pause au milieu de la phrase 1bis de cette manière : « Elle manqua / mon amie » pour être sûrs de la distinguer de la première « Elle manque à mon amie ». Les VTO et les T n'auraient pas eu besoin de cette pause, ce qui expliquerait une très faible différence entre le nombre de pauses de chaque phrase, observée en moyenne pour ces deux groupes. Le couple 4 montre une faible différence du nombre de pauses entre les deux phrases pour les groupes VO et VTO et une différence importante pour les sujets T. La phrase 4 étant une des plus longues et complexes à prononcer, il est possible que les patients aient eu du mal à respecter les pauses qu'on aurait attendu (qui sont respectées par le groupe T). Les locuteurs ne font pas toujours de pause. Pour le couple 5 par exemple, la plupart des sujets a fait une pause à la phrase 5 et pas à la phrase 5bis, ce qui nous fait déduire que la pause est, pour ce couple (n°5), un critère majeur de distinction. *A contrario*, les pauses faites pour la phrase 4 « Je suis inquiet de la varice de ma mère » et 4bis « Je suis inquiet de l'avarice de ma mère » varient beaucoup d'un sujet à un autre, même au sein du groupe témoin. Il doit donc y avoir pour ce couple (n°4), d'autres tendances temporelles identifiables qu'il serait intéressant de rechercher (allongement des voyelles, des syllabes et autres).

Annexe B

Exemples de segmentation à partir de l'enregistrement de la lecture des phrases ambiguës du couple 5 par un sujet VO. Nous voyons du haut vers le bas : le signal acoustique, le spectrogramme, la ligne de segmentation de mots orthographiés, la ligne de segmentation de phonétique des mots et la ligne de segmentation des phonèmes. Nous pouvons voir que le sujet VO a fait une pause entre « jean » et « mène » (indiquée par le signe #) pour la phrase « Jean mène un journal » (image du haut) et n'a pas fait de pauses pour la phrase « J'emmène un journal ».

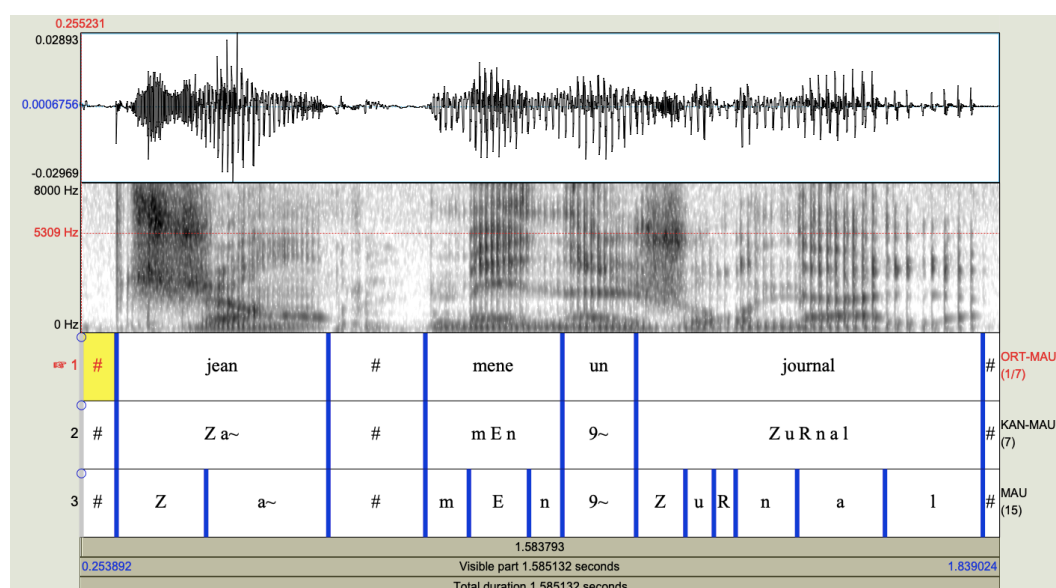


Figure 6 : Segmentation via Praat de la phrase « Jean mène un journal », lue par un sujet VO.

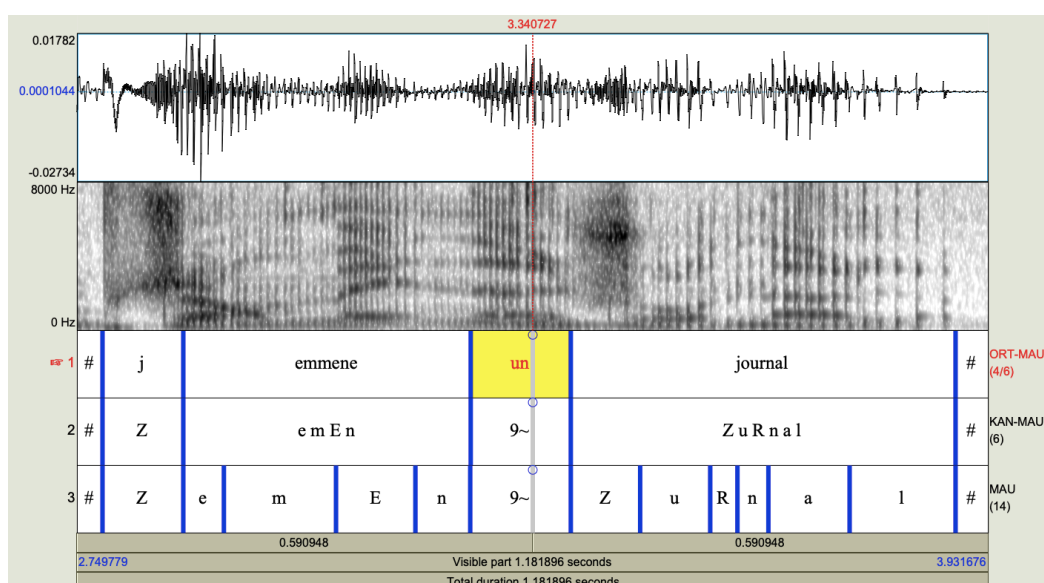


Figure 7 : Segmentation via Praat de la phrase « Jean mène un journal », lue par un sujet VO.

Annexe C

Tableau 2. Scores bruts du VHI (Physique /40, Fonctionnel /40, Emotionnel /40 et Total /120) du groupe VO et du groupe VTO.

Groupe	Physique	Fonctionnel	Emotionnel	Total
VO Moyenne : 41,2 Ecart-type : 22	13	8	3	24
	15	4	2	11
	15	12	4	31
	7	14	4	25
	17	18	1	36
	23	14	21	58
	18	12	9	39
	24	21	20	65
	9	25	2	36
	29	24	17	70
	12	2	5	19
	27	30	23	80
	37	9	9	37
VTO Moyenne : 33,6 Ecart-type : 22	8	4	1	13
	10	8	2	20
	32	20	26	78
	5	8	3	16
	3	6	3	12
	16	16	9	41
	11	17	12	40
	6	15	1	22
	22	11	9	42
	0	10	1	11
	11	12	9	32
	27	30	23	80
	12	10	5	27

Annexe D

Tableau 3. Moyennes - Dimensions EORTC QLQ H&N 35. Scores de fréquence /4 soit 1 pour « pas du tout », 2 pour « peu », 3 pour « beaucoup » 4 pour « toujours ».

	VO	VTO
Déglutition	1,2	1,73
Douleur	1,14	1,17
Sens	1,2	2,64
Parole	1,6	1,43
Mâchoire	1,4	1,68
Alimentation public	1,6	1,46
Contact social	1,5	1,21
Altération sexuelle	1,5	1,27
Souci	1,2	1,32
Moyenne / Ecart-type	1,37 / 0,187	1,86 / 0,454

Résumé

Cette étude dresse un état des lieux de la parole après laryngectomie totale (LT) en se focalisant sur les paramètres temporels globaux (d'un énoncé) et locaux (à partir de segments de parole). Un corpus composé d'items issus de la littérature a permis d'évaluer les différents paramètres temporels dans trois situations (lecture, discours spontané et diadococinésies). Il a été soumis à une population composée de 12 sujets utilisant la voix oro-œsophagienne (VO), 17 la voix trachéo-œsophagienne (VT) et de 12 sujets témoins (T). Les résultats ont montré que la vitesse de parole des patients, en particulier de ceux utilisant la VO, était ralentie, dans toutes les situations. La fréquence des pauses et le nombre d'expirations élevés démontrent aussi un impact réel de la LT sur la rapidité et la fluidité de l'élocution. Des mesures plus précises autour des voyelles ont été réalisées : de légères différences entre les groupes VO et T et un rapprochement entre les groupes VTO et T sont observables. Les perturbations de la parole peuvent troubler la qualité de vie et sont en lien avec la voix. Des corrélations significatives positives ont été observées entre des items de qualité de vie et de handicap vocal ; de plus, les sujets LT ressentiraient un handicap vocal plus élevé lorsqu'ils présentent une vitesse de lecture basse et inversement.

The present study is a state of the art of speech ability after total laryngectomy. It focus on global (statement) and local (speech segments) temporal parameters of elocution. A text corpus based on literature extracts enabled to evaluate various temporal parameters through three categories (reading, spontaneous speech and diadokokinesis). The corpus was submitted to (1) a 12-subjects population using esophageal voice, (2) a 17-persons population using tracheo-esophageal voice and (3) a 12-subjects control group. Results show that patients' speech velocity was reduced for the three category of communication, in particular for the patients with esophageal voice. Pauses frequency as well as the expiration number also show a significant impact of total laryngectomy on elocution velocity and fluidity. More precise measurements were performed regarding vowels : minor difference were observed between esophageal and control groups whereas similar results were obtained between tracheo-esophageal and control groups. Speech malfunction might trouble life quality and are linked to voice quality and voice handicap. Significant correlations were observed between certain aspects of life quality and voice handicap ; moreover laryngectomised subjects were shown to feel a strong voice handicap when presenting a slow read velocity and inversely.