Modifications rythmiques de la parole induites par la laryngectomie totale

Camille Doumayrou, Marie Drieux

▶ **To cite this version:**

Camille Doumayrou, Marie Drieux. Modifications rythmiques de la parole induites par la laryngectomie totale. Sciences du Vivant [q-bio]. 2021. dumas-03349356

HAL Id: dumas-03349356
https://dumas.ccsd.cnrs.fr/dumas-03349356

Submitted on 20 Sep 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.

Faculté des sciences médicales et paramédicales
Aix*Marseille Université

Centre de Formation Universitaire en Orthophonie de Marseille

Modifications rythmiques de la parole induites par la laryngectomie totale

Mémoire de fin d'études en vue de l'obtention du Certificat de Capacité d'Orthophoniste

Camille DOUMAYROU et Marie DRIEUX

Sous la direction de Camille GALANT

14 juin 2021

UMR 7309 • CNRS • AMU
PAROLE ET LANGAGE

Remerciements

Nous tenons à remercier, pour leur contribution à l'élaboration de notre mémoire :

Camille GALANT, notre maître de mémoire, pour sa confiance, son soutien et sa motivation sans faille. Ce travail n'aurait pas été possible sans son implication démesurée et ses conseils constructifs s'inscrivant dans une réelle démarche collaborative. Nous tenions aussi à souligner son dévouement pédagogique tout au long de notre formation que ce soit de par la qualité de ses enseignements ou de par son implication active dans le bon fonctionnement de notre parcours étudiant.

Alain GHIO pour son aide précieuse dans l'analyse statistique de nos résultats et pour sa disponibilité.

Muriel LALAIN, pour avoir accepté de faire partie de notre jury de soutenance.

Nous voulons aussi remercier chaleureusement tous nos enseignants qui ont forgé ces 5 années d'études, et plus particulièrement les orthophonistes qui ont su nous guider vers ce métier si complet et si complexe.

Je remercie ma famille et mon petit ami qui ont été le soutien rêvé pour le bon déroulement de mon choix de parcours, de mes études et de ce mémoire. Sans leurs encouragements, leur écoute, leur disponibilité et leur amour je n'en serais sûrement pas là aujourd'hui.
Mes amis, et tout particulièrement Mathilde qui a été la petite lumière de positivité et de joie de vivre dans ces 5 années parfois compliquées. Je ne pensais pas rencontrer un jour la gentillesse incarnée, reste comme tu es.
Et je conclurai bien évidemment par toi Camille, mon binôme de cœur. Je n'ai pas assez d'espace pour te dire tout ce que j'aurai à te dire et je suis bien trop pudique, mais sache que je suis très heureuse d'avoir partagé ce parcours du combattant avec toi. Tu as toujours répondu présent que ce soit pour me motiver, me faire rire, ou m'apaiser dans des situations où je me noyais dans un verre d'eau. Nous sommes si différentes et si semblables à la fois que nous avons fait ressortir le meilleur l'une de l'autre. Ces cinq ans resteront un merveilleux souvenir pour moi. Tu es douce, compréhensive, très empathique et je suis sûre que tu seras une merveilleuse orthophoniste.

Marie

Je remercie mes parents, qui m'ont soutenue à tous les niveaux, à la fois durant ce travail de mémoire, mais également tout au long de ma formation. Sans vous, je n'aurais pas eu la possibilité de faire les études menant à ce beau métier. Ma sœur, bien sûr, qui a toujours su trouver les mots lorsque je perdais confiance, et a été un soutien sans faille. Je vous aime.
Je tiens également à dire merci à mes amis, d'ici et d'ailleurs, et spécialement à notre amie Mathilde, qui nous est chère. Nous avons partagé cinq années, avec des moments de doutes

dans lesquels tu as toujours su nous redonner le sourire. J'espère que notre amitié perdurera et que nos chemins, professionnels et privés se croiseront souvent.
Enfin, celle qui a partagé ces deux années de travail, mais aussi aujourd'hui cinq ans de ma vie, Marie. Tu es une fille extraordinaire, pleine de ressources, appliquée dans son travail et toujours encline à tendre la main. Sans toi, j'aurais plus d'une fois perdu pied. Tu as été la lanterne qui a éclairé ce chemin semé d'embûches, et je suis plus qu'heureuse de l'avoir parcouru avec toi. Nous nous sommes complétées, écoutées et épaulées, afin de produire le meilleur travail possible. Je suis sûre que tu seras la meilleure des orthophonistes, tu as toutes les qualités requises, et j'espère du plus profond de mon cœur que notre amitié restera telle quelle.

Camille

Table des matières

Introduction ____ *5*

1. Etat des lieux ____ *7*

1.1. Généralités ____ **7**

1.2. Evaluation de la voix oro-œsophagienne ____ **8**

1.2.1. Mesures acoustiques et aérodynamiques ____ 8

1.2.2. Evaluation perceptive ____ 10

1.2.3. Auto-évaluation ____ 12

1.2.4. Conclusion sur les évaluations ____ 13

1.3. Parole et rythme ____ **13**

1.3.1. Rythme et pauses dans la parole normale ____ 14

1.3.2. Perturbation du rythme de la parole ____ 17

1.3.3. Rythme et pauses dans la parole oro-œsophagienne ____ 17

2. Méthodologie de la recherche ____ *20*

2.1. Expérimentation ____ **20**

2.1.1. Population ____ 20

2.1.2. Corpus ____ 20

2.2. Traitement des données ____ **21**

2.2.1. Transcription orthographique ____ 21

2.2.2. Segmentations ____ 22

2.2.3. Etiquetage des pauses ____ 23

2.2.4. Extraction des données ____ 25

2.3. Analyses statistiques ____ **25**

3. Résultats ____ *26*

3.1. Analyse des pauses ____ **26**

3.1.1. Quantité ____ 26

3.1.2. Durée ____ 31

3.2 – Durée des syllabes ____ **36**

3.2.1 - Comparaison de l'allongement final aux frontières prosodiques ____ 36

3.3. Débit de parole ____ **38**

3.4. Vitesse articulatoire ____ **38**

4. Discussion ____ *40*

5. Limites et perspectives ____ *48*

6. Conclusion ____ *50*

Bibliographie ____ *52*

Annexes ____ *58*

Résumé ____ *67*

Introduction

Le larynx, organe situé au creux de la gorge, est souvent méconnu, voire négligé. Or, celui-ci est essentiel à différentes fonctions vitales, telles que la respiration ou la déglutition. On ne se rend pourtant compte du trésor qu'il contient – notre voix – que lorsque nous le perdons. Organe assurant la fonction de phonation, c'est bien lui qui permet à l'homme de communiquer par le biais du langage, possible grâce au son créé par cet instrument. D'un point de vue anatomique, il s'agit d'un ensemble d'os, de cartilages, sur lesquels viennent se fixer les plis vocaux, couramment appelés "cordes vocales". Ces muscles, dont nous nous servons inconsciemment chaque jour, nous les avons travaillés, modelés, pour qu'ils portent à travers notre voix notre personnalité, nos émotions, les messages que nous voulons faire passer à l'autre. Ils participent donc pleinement à établir notre identité.

La laryngectomie totale est une opération visant à guérir un cancer étendu de la gorge, consistant en l'ablation de cet organe. Sans lui, nous perdons alors notre instrument de voix, de communication. Il faut alors retrouver un moyen de nous exprimer, un moyen de communiquer, et cela passera par l'apprentissage, l'appropriation d'une nouvelle voix. L'orthophoniste joue un rôle majeur dans l'accompagnement de ces patients néo-laryngectomisés en participant pleinement à leur réhabilitation. Selon les définitions de l'OMS, la réhabilitation contient trois objectifs principaux : 1) la rééducation visant à réduire les déficiences et incapacités du patient (ici, la rééducation vocale à travers l'apprentissage d'une voix de substitution), 2) la réinsertion qui comprend l'ensemble des mesures ayant pour objectif le retour en société et éviter l'exclusion, 3) la réadaptation correspondant aux différents moyens mis en œuvre pour aider le patient à s'adapter à ses incapacités (*World Report on Disability*, 2011). L'apprentissage de la voix oro-œsophagienne rentre donc majoritairement dans l'objectif de rééducation, mais participe aussi à la réadaptation du patient et à sa réinsertion dans la société. La voix oro-œsophagienne constitue l'un des moyens de communication orale après LT. Cette voix dite de "substitution" provient de l'alternance de blocages d'air dans l'œsophage et d'éructations contrôlées. Ces dernières font vibrer la muqueuse néo-pharyngée, remplaçant celle des cordes vocales. Ce nouveau mécanisme de phonation entraîne de nombreuses modifications : la fréquence fondamentale qui caractérisait notre voix est plus basse, voire irrégulière ; l'intensité, la stabilité et la qualité vocale sont altérées ; la durée des rhèses est plus courte du fait d'un réservoir d'air œsophagien plus restreint, etc...

L'esthétique de la voix est donc changée, ce qui est d'ailleurs ressenti en premier plan par ces patients. Toutefois, ils doivent parler sans articulateur laryngé tout en essayant de rester intelligibles pour leur interlocuteur, et cela entraîne donc également des modifications inévitables au niveau de leur parole. Le patient utilisant une voix oro-œsophagienne doit bloquer de l'air pour pouvoir produire du son, cela implique donc l'adjonction de nombreuses pauses dans son discours. De fait, la prosodie de la parole est altérée : au niveau du rythme, car le discours est dorénavant entrecoupé de pauses ; mais aussi au niveau mélodique, car le patient ne peut contrôler la vibration phonatoire comme il le faisait avec ses cordes vocales.

Si la voix oro-œsophagienne a été largement décrite dans la littérature, nous avons toutefois observé que la parole n'était quant à elle que très peu étudiée. Pourtant, il semble

que la prosodie, et particulièrement le rythme, joue un rôle essentiel dans la transmission d'un message. Le but de ce mémoire sera dans un premier temps d'établir une description de l'organisation temporelle de la parole du sujet utilisant une voix oro-œsophagienne, afin de déterminer si l'axe rythmique de la parole pourrait devenir un paramètre objectif dans l'évaluation de cette voix, et par extension un enjeu de sa prise en charge.

1. Etat des lieux

1.1. Généralités

Selon l'INCa, en 2018, on estime à 3 160 le nombre de nouveaux cas de cancer du larynx en France métropolitaine, dont 87% sont des sujets masculins. Les facteurs de risque les plus courants sont la consommation de tabac et d'alcool. L'âge moyen lors du diagnostic est de 64 ans chez l'homme et 62 ans chez la femme. Chez l'homme, l'incidence des cancers du larynx a nettement diminué entre 1990 et 2018, une tendance inverse est observée chez la femme (Defossez G, et al, 2019).

La laryngectomie totale est une opération chirurgicale indiquée face à des cancers du larynx de stade avancé (T4 = envahissement cartilagineux) ou face à une impossibilité ou un échec de protocole de préservation d'organe (T3 avec contre-indication à la radio/chimiothérapie ou laryngectomie totale de rattrapage) (Antin, 2017). Elle consiste en l'ablation du bloc laryngé, ce qui entraîne le détournement complet des voies respiratoires vers le cou, la création d'un trachéostome (Graville & Palmer, 2018) et la perte de la phonation laryngée. Le larynx pouvant être considéré comme l'articulateur de la parole, le patient opéré d'une laryngectomie totale perdra définitivement sa voix et sa parole d'origine (Galant, 2019).

Cette opération est une mutilation vécue comme un véritable traumatisme par les patients laryngectomisés, car elle engendre de nombreuses séquelles anatomiques, physiologiques, psychologiques et relationnelles. La personne laryngectomisée peut être considérée comme un *mutilé de la voix* : elle a perdu cette capacité vocale à laquelle on attribue de nombreuses fonctions comme celles de communication verbale et émotionnelle, d'outil social, de représentation de sa personnalité et d'affirmation de son statut d'individu. En effet, la voix est considérée comme une carte d'identité pour chaque individu (Bretagne, 2008; Harrus-Révidi, 2007).

La réhabilitation vocale devient alors un objectif majeur de la rééducation de l'individu ayant bénéficié d'une laryngectomie totale. La principale direction que prend l'intervention orthophonique est donc l'apprentissage et la maîtrise d'une voix de substitution, définie par une phonation produite sans plis vocaux (Moerman et al., 2004). Il existe différents modes de phonation post-laryngectomie dont les principaux sont : la voix oro-œsophagienne (VOO), la voix trachéo-œsophagienne (VTO) et celle obtenue via un électrolarynx (van Sluis et al., 2018). Les deux premières ont pour point commun d'avoir une source de phonation interne, produite via le segment pharyngo-œsophagien (PE) ou néo-glotte, mais des sources d'air différentes (VOO source d'air œsophagienne / VTO source d'air pulmonaire) (Ng, 2019), tandis que la troisième implique une source de phonation externe, l'électrolarynx (van Sluis et al., 2018). Dans notre travail, nous nous intéresserons plus particulièrement à la voix oro-œsophagienne.

La voix oro-œsophagienne repose sur la production d'une éructation contrôlée, obtenue grâce à un blocage d'air dans l'œsophage puis une expulsion volontaire de celui-ci faisant vibrer la muqueuse néopharyngée (Zenga et al., 2018). La parole réalisée au moyen de la phonation alaryngée s'appelle l'érygmophonie (Allali et al., 2008; Harrus-Révidi, 2007). Son

apprentissage est long et laborieux. Il nécessite beaucoup de motivation, de persévérance et une participation active du patient, mais la maîtrise finale de cette voix n'est jamais garantie (Bretagne, 2008; van Sluis et al., 2018). En effet, seulement 2/3 des laryngectomisés obtiennent une voix oro-œsophagienne fluide (Zenga et al., 2018). De plus, l'éructation volontaire engendre une réticence de la part du patient du fait de son inesthétisme (Bretagne, 2008; Harrus-Révidi, 2007) et de sa faible acceptabilité sociale.

1.2. Evaluation de la voix oro-œsophagienne

Les sujets opérés par laryngectomie totale font donc face à plusieurs obstacles. Le plus explicite est la dégradation de leur voix. Ils doivent faire le deuil de leur voix naturelle et apprendre à sonoriser autrement (Dragičević et al., 2019; Galant, 2019). L'érygmophonie nécessite de produire des phonèmes sans articulateur laryngé tout en s'assurant que la parole reste intelligible (Galant, 2019).

1.2.1. Mesures acoustiques et aérodynamiques

La voix alaryngée étant produite par la vibration de la néoglotte, la qualité de la voix est lourdement impactée et diffère fondamentalement de celle produite par la vibration des cordes vocales. Cela peut s'évaluer par la perturbation de divers paramètres acoustiques largement étudiés dans la littérature :

- *La fréquence fondamentale ou F0 :* Des résultats hétéroclites ont été obtenus par différentes études au sujet de la fréquence fondamentale de la VOO. Certains auteurs observent une fréquence fondamentale plus basse chez le sujet oro-œsophagien que chez le sujet sain (Bellandese et al., 2001; Globlek et al., 2004; Most et al., 2000), d'autres observent en cantonais une F0 plus haute pour la VOO (Ng et al., 2001; Yan et al., 2012) qui serait due à un meilleur contrôle de la néoglotte dans les langues à ton. Enfin, certains ne remarquent pas de différence significative entre VOO et sujet sain (Moerman et al., 2004). De la même manière, pour certains auteurs, peu de différences sont retrouvées entre la fréquence fondamentale en VOO et celle en VTO (Bellandese et al., 2001; Most et al., 2000; Yan et al., 2012), probablement en lien avec le fait qu'ils aient la même source sonore, alors que pour d'autres la fréquence fondamentale en VOO est plus basse qu'en VTO (Allegra et al., 2019; van Sluis et al., 2018). Cette fréquence fondamentale permet la distinction de genre en voix laryngée. Ceci a été confirmé par Bellandese pour les locuteurs féminins en VOO et VTO, mais pas pour les locuteurs masculins (Bellandese, 2009). De plus, van Rossum retrouve dans son travail de thèse des disparités en ce qui concerne la F0 chez les sujets oro-œsophagiens : parfois la F0 est présente mais erratique, parfois elle est absente. Les résultats sont divers et la mesure significative de ce paramètre est difficile (van Rossum, 2005). Ces différentes réflexions et cette disparité nous amènent à questionner l'intérêt de cette mesure dans l'évaluation de la VOO.

- *Volume/intensité* : L'intensité vocale est considérée par la majorité des auteurs comme plus faible chez le sujet oro-œsophagien que chez le sujet sain (Globlek et al., 2004; Ng et al., 2001; Ng, 2019). Cette intensité est influencée par le réservoir d'air et donc la pression

exercée sur le segment pharyngo-œsophagien. Cela est prouvé par le fait que les VTO ont une meilleure intensité que les VOO (Globlek et al., 2004; Ng, 2019; van Sluis et al., 2018). Seule l'étude publiée par Motta en 2001 ne trouve pas de différence significative entre l'intensité de la VOO et celle de la VTO (Motta et al., 2001). Cette intensité ne serait d'ailleurs pas corrélée à l'intelligibilité de la VOO (Ng et al., 2001). De la même manière, dans son étude de l'effet Lombard sur les voix de substitutions, Ng a démontré que les sujets oro-œsophagiens pouvaient augmenter le volume de leur voix en cas de bruit de fond, mais que cela n'améliore pas leur intelligibilité (Ng, 2019).

- *Jitter et shimmer* : Le jitter est un indice de la stabilité vocale et le shimmer évalue la variation d'intensité. Ces deux paramètres acoustiques temporels sont plus importants chez le sujet oro-œsophagien que chez le sujet sain ou le sujet VTO (Globlek et al., 2004; MacCallum et al., 2009; Moerman et al., 2004; van Sluis et al., 2018). Mais ces mesures sont considérées dans l'ensemble des études comme très peu fiables dans l'évaluation des voix de substitution du fait de l'apériodicité de leur signal acoustique.

- *Rapport signal/bruit ou HNR* : Le rapport harmonique sur bruit mesure la qualité vocale, il révèle le souffle de la voix. Comme on pouvait s'y attendre, ce rapport est perturbé chez les sujets utilisant une VOO. En effet, le HNR est plus bas que chez les sujets sains (Bellandese et al., 2001 ; MacCallum et al., 2009) et que chez les sujets VTO (Allegra et al., 2019; van Sluis et al., 2018) et est décrit comme peu pertinent dans l'évaluation de la VOO.

- *Voisement* : La mise en vibration du segment pharyngo-œsophagien est plus laborieuse que celle des cordes vocales et est d'autant plus difficile à gérer. La qualité de voisement est donc impactée chez les patients utilisant la VOO. En 2004, Moerman rapporte chez le sujet VOO plus de difficultés à réaliser des voisements longs dans un contexte voyelle-consonne et, dans une autre étude en 2015 sur l'analyse de la VOO grâce à l'analyse acoustique AMPEX et l'échelle perceptive INFVo, elle observe que le voisement impacte l'intelligibilité ainsi que l'impression de qualité. Il serait de plus un paramètre très informatif dans l'analyse acoustique (Moerman et al., 2004; Moerman et al., 2015). Most, dans son étude sur les caractéristiques de la VOO et VTO, retrouve une amplitude des voyelles plus élevée chez les locuteurs VOO moyens que chez les bons locuteurs VOO. Elle l'associe à la qualité de gestion du voisement, obtenue grâce au contrôle de l'utilisation de l'air et du relâchement du segment PE (Most et al., 2000).

On observe donc que, dans la littérature, certains paramètres acoustiques sont relativement stables et d'autres ont des résultats plus divers. Du fait de l'apériodicité du signal acoustique de la voix oro-œsophagienne, les paramètres acoustiques sont peu fiables et décrits comme peu adaptés à l'évaluation de la VOO (MacCallum et al., 2009; Moerman et al., 2004; Yan et al., 2013). Moerman, ayant utilisé le système AMPEX (Auditory Model Based Pitch Extractor) dans une étude publiée en 2015, a pu confirmer l'impertinence des mesures fondées sur la perturbation du signal sonore pour évaluer objectivement les voix alaryngées (Moerman et al., 2015).

De la même manière, certains paramètres aérodynamiques seraient impactés dans l'utilisation de la voix alaryngée :

Concernant le temps maximal de phonation (TMP), on retrouve dans l'ensemble des études que les sujets VTO ont un meilleur TMP que les sujets VOO, et qu'il existe même une différence entre bon et mauvais locuteurs VOO : les bons locuteurs auraient un TMP plus long. Ceci serait principalement dû à la capacité du réservoir d'air et à la gestion du relâchement du segment PE (Globlek et al., 2004; Motta et al., 2001; van Sluis et al., 2018).

De plus, Motta, dans son étude des paramètres aérodynamiques des bons et mauvais locuteurs VOO et des locuteurs VTO en 2001, a observé que les sujets ayant une VOO médiocre produisent une hyperarticulation consonantique (augmentation de la pression intra-orale) dans un objectif d'intelligibilité, mais que cela altère d'autant plus leurs productions. De plus, malgré leur réserve d'air phonatoire faible, ils présentent un débit phonatoire élevé et un débit résiduel post-phonation plus important que les autres groupes, ce qui engendre un TMP réduit, l'ajout de nombreuses pauses et donc l'altération de leur intelligibilité. Sachant les difficultés de coordination pneumo-phonique inerrantes à la laryngectomie totale, cette augmentation de pression peut engendrer un souffle trachéal important au cours de la phonation. Ce dernier est d'ailleurs à l'origine d'une plainte régulière chez le sujet oro-œsophagien (Motta et al., 2001).

Les mesures aérodynamiques apparaissent plus fiables que les mesures acoustiques mais sont plus contraignantes à évaluer dans un contexte clinique car nécessitent un matériel peu fréquent et onéreux.

1.2.2. Evaluation perceptive

Comme nous l'avons vu précédemment, les modifications induites par la laryngectomie totale engendrent une dysphonie sévère avec des perturbations de la majorité des mesures acoustiques et aérodynamiques. L'évaluation subjective des voix de substitution a pour objectif d'évaluer la perception de sévérité de l'atteinte vocale. Cette méthode a l'avantage d'évaluer l'expression sous sa forme « écologique ». En effet, la perception qu'ont les interlocuteurs de la voix et de la parole du patient est une des plaintes majeures des personnes utilisant une voix de substitution. L'objectif final de la rééducation orthophonique est d'obtenir une communication efficace, ce qui inclut les productions du patient mais aussi la perception qu'en ont ses partenaires de communication.

Si l'évaluation perceptive est facile d'application, elle comprend toutefois les nombreux inconvénients de la subjectivité : le jugement de la voix diffère d'une personne à l'autre et peut aussi varier chez un même juge, c'est ce qu'on appelle les variabilités inter et intra-individuelles. Il est donc nécessaire de minorer cette subjectivité autant que possible. Pour cela, Hurren et al. ont étudié les variabilités inter et intra-individuelles des juges et des conditions ayant des conséquences sur l'évaluation perceptive via l'échelle SToPS (Sunderland Tracheoesophageal Perceptual Scale). Ils ont démontré que la formation à l'analyse perceptuelle et le niveau d'expertise de l'évaluateur avaient un impact important sur l'évaluation perceptive, et que la fiabilité intra et inter-évaluateurs était supérieure pour les paramètres globaux de type « gravité globale » (note globale et acceptabilité sociale) par rapport à d'autres qui nécessitent une discrimination plus complexe (paramètres « unidimensionnels ») (Hurren et al., 2019). Des conclusions similaires ont été retrouvées dans

l'étude de Suhail et al. et ils ont recommandé que cette évaluation soit effectuée par des évaluateurs expérimentés, car les évaluateurs naïfs auraient tendance à juger ces voix aux extrémités inférieures des échelles. (Suhail et al., 2016 in Dragičević et al., 2019).

Le comité de phoniatrie de la Société Européenne de Laryngologie (ELS) recommande l'évaluation perceptive comme l'un des cinq éléments essentiels de l'évaluation de la voix (perception, vidéo-stroboscopie, analyse acoustique, mesures aérodynamiques et auto-évaluation perceptive). Le groupe a préconisé l'utilisation de l'échelle GRBAS (I) comme méthode la plus pratique pour évaluer la perception (Dejonckere et al., 2001). Celle-ci permet d'évaluer plusieurs paramètres de la voix : appréciation globale, éraillement, présence d'une fuite d'air, hypotonie ou hypertonie vocale, en les notant de 0 (normal) à 3 (très altéré). Cependant, il semble que ceux-ci n'apportent pas de résultats significatifs dans le cas de l'évaluation des voix de substitutions, car leur sévérité engendre des scores saturés sans possibilité d'analyse plus précise. Moerman et al. ont validé en 2006 une échelle d'évaluation perceptive spécifique aux voix de substitution appelée (I)INFVo. Elle propose 5 nouveaux paramètres d'évaluation (impression globale, intelligibilité, bruit parasite, fluidité de la parole et qualité de voisement) cotés sur une échelle visuelle analogique de 0 (très bonne voix de substitution) à 10 (mauvaise voix de substitution). Dans leur étude, ils ont aussi démontré la redondance des deux paramètres « I » car l'intelligibilité participe à l'impression globale de qualité vocale, créant alors l'échelle INFVo démontrée comme adaptée et opérante pour les voix de substitution (Moerman et al., 2006, 2015).

Différents auteurs ont étudié certains paramètres perceptuels. Motta en 2001 a observé que le souffle trachéal, bruit parasite causé par une mauvaise coordination des mécanismes phono-respiratoires, et le nombre de pauses phonatoires (atteinte de la fluidité) différencient statistiquement la VTO de la VOO et les bons locuteurs VOO des mauvais (Motta et al., 2001).

Pour ce qui est de l'intelligibilité, Ng décrit en 2009 que la parole alaryngée est associée à une intelligibilité diminuée en raison d'une détérioration de la qualité de la voix et d'un manque de variabilité de la hauteur et de l'intonation. La réduction de la force et de la mobilité linguales provenant de l'excision de l'os hyoïde pendant la laryngectomie totale pourrait compromettre la précision articulatoire et le rythme (Ng, 2019). Motta en 2001 évoque un lien entre nombre de pauses et intelligibilité, avec ici aussi de meilleurs résultats pour la VTO que pour la VOO et pour les bons VOO que pour les mauvais. Mais dans cette étude la mesure des pauses avait eu lieu sur une épreuve de comptage, ce qui représente un biais important car il s'agit d'une modalité non représentative de la parole spontanée, pouvant induire un rythme de la parole saccadé (Motta et al., 2001).

Most, en 2000, différencie intelligibilité (clarté, identification et compréhension des mots et phonèmes) et acceptabilité (impression subjective du discours et de la voix concernant sa qualité et sa fluidité, ses caractéristiques prosodiques et son articulation). Dans les deux cas les locuteurs oro-œsophagiens étaient inférieurs aux locuteurs laryngés et trachéo-œsophagiens (corrélations élevées entre ces deux critères). Des corrélations significatives ont été observées entre l'acceptabilité et certaines mesures acoustiques (syllabes par seconde, durée de la phrase, durée de la syllabe et la F0) et l'intelligibilité des mots et des phonèmes était fortement corrélée à la fréquence fondamentale (Most et al., 2000). De la même

manière, Bellandese retrouve en 2001 des résultats sur les variables acoustiques en lien avec l'acceptabilité (F0, syllabes par minute, nombre moyen de pauses, etc…) significativement différents entre les deux groupes alaryngés (excellentes locutrices VTO et VOO) et le groupe laryngé (Bellandese et al., 2001)

Enfin Moerman et al., en 2015, observent dans leur étude sur l'évaluation acoustique AMPEX (the Auditory Model Based Pitch Extractor) que la fluidité de production vocale et le voisement déterminent l'intelligibilité perçue. Or, ces deux paramètres sont très dépendants des pauses : la longueur de voisement (paramètre *VL90*) était définie comme le nombre de trames voisées consécutives trouvées dans les données (les pauses phonatoires réduisent ce paramètre). La qualité du voisement était objectivée grâce au rapport PVF (proportion de trames voisées) / PVS (proportion de trames voisées attendues), qui est influencé par le nombre de pauses (PVF dépend de ces pauses et des sons faibles non voisés). Enfin, la fluidité est bien sûr impactée par toute perturbation du rythme (Moerman et al., 2015).

De plus, dans leur étude sur l'évaluation multidimensionnelle, Moerman et al. ne retrouvent que peu ou pas de corrélation entre l'auto-évaluation, l'évaluation acoustique et l'évaluation perceptive. Il est donc nécessaire d'avoir une approche multidimensionnelle pratique et peu chronophage dans l'évaluation vocale des voix fortement irrégulières. Il existerait une très faible corrélation entre la perception du clinicien et la perception de soi du patient, démontrant donc l'importance des auto-évaluations (Moerman et al., 2015).

1.2.3. Auto-évaluation

La voix fait partie intégrante de notre identité personnelle, il est donc primordial d'obtenir le ressenti de la personne présentant un trouble vocal. L'auto-évaluation a été développée dans ce sens et a aujourd'hui une place majeure dans l'évaluation de la communication et l'adaptation des objectifs thérapeutiques dans le but d'une amélioration de la qualité de vie. Le gold standard actuel est le VHI, ou Voice Handicap Index, développé par Jacobson et al. (Jacobson Barbara H. et al., 1997). Cette échelle d'évaluation de trente items comporte trois sous-échelles (émotionnelle, fonctionnelle, physique) permettant d'évaluer le handicap vocal au quotidien et de préciser les domaines le plus atteints.

Le VHI a été largement étudié dans la littérature pour ce qui est de la voix saine et des pathologies vocales laryngées, mais il a aussi été utilisé dans des études sur l'auto-évaluation des voix alaryngées. Contrairement à ce que l'on pouvait penser, l'auto-évaluation des patients utilisant une voix oro-œsophagienne n'est pas très sévère : l'incapacité due au trouble de leur voix est considéré comme modéré. Cela pourrait s'expliquer par une certaine tolérance envers leur voix, au vu des efforts fournis pour l'obtenir et de l'enjeu vital de leur cancer (Moerman et al., 2004; van Sluis et al., 2018).

Allegra observe aussi qu'un niveau d'éducation plus élevé et une plus grande distance à la laryngectomie sont corrélés à une meilleure évaluation subjective de leur performance verbale (Allegra et al., 2019). En comparant VOO et VTO, le handicap vocal serait tout de même plus important chez les sujets oro-œsophagien que chez les sujets trachéo-œsophagien. Les différences majeures se retrouvent pour la sous-échelle fonctionnelle d'après Allegra et al. dans leur étude sur différentes échelles d'auto-évaluations (VHI, V-RQOL

- *Voicerelated Quality of Life* et VQP - *Voice Performance Questionnaire)*, mais Dragičević et al. retrouvent des différences dans les trois sous-échelles du VHI (Allegra et al., 2019; Dragičević et al., 2019). Même si la majorité des résultats montrent des niveaux différents de handicap vocal, aucune supériorité significative n'a pu être démontrée entre les méthodes de réhabilitation vocale (Allegra et al., 2019; Dragičević et al., 2019; van Sluis et al., 2018).

D'après Moerman et al., le VHI serait un outil adéquat pour évaluer la gravité des troubles vocaux dans les voix de substitution. Ils observent une large répartition des scores, ce qui montre que l'échelle n'est pas saturée vers les valeurs extrêmes comme attendu dans des pathologies aussi sévères. Mais il serait plus adapté à l'évaluation d'un même patient au cours du temps qu'à une comparaison entre pathologies diverses. Dans leur étude, ils ont adapté ce questionnaire en laissant les patients ne pas répondre à des items qu'ils considéraient comme non pertinents vis-à-vis de leur situation actuelle pour ne pas induire des choix forcés qui perturberaient les résultats. Les domaines physiques et fonctionnels contiendraient plus d'items non ou peu pertinents que le domaine émotionnel (Moerman et al., 2004). Les résultats au VHI ne sont pas non plus influencés de manière significative par l'âge ni par l'irradiation (Dragičević et al., 2019; Moerman et al., 2004).

De plus, il existerait une variation conséquente des résultats entre les sujets. L'écart-type important dans les groupes de sujets laryngectomisés reflète une grande variation inter-individuelle de l'auto-évaluation (Dragičević et al., 2019; van Sluis et al., 2018). Chaque individu est unique du fait de sa personnalité, ses besoins, ses attentes et sa vie relationnelle. Le handicap induit par la laryngectomie totale n'impacte donc pas tout le monde de la même manière, d'où l'importance de l'auto-évaluation qui apporte des informations capitales de la part de l'acteur principal de cette réhabilitation : le patient lui-même. Il est clair aujourd'hui que les évaluations objectives et subjectives réalisées par un examinateur ne suffisent pas à obtenir une approche globale et pertinente de la situation de la personne laryngectomisée.

1.2.4. Conclusion sur les évaluations

D'après Moerman et al., la corrélation entre les paramètres subjectifs et objectifs de l'évaluation de la voix oro-œsophagienne n'est actuellement pas satisfaisante. Pour atteindre une échelle objective utilisable pour les voix de substitution, il sera nécessaire d'explorer différentes combinaisons de paramètres, mais aussi d'en rechercher de plus objectifs pour caractériser ces voix et en obtenir une évaluation efficace et pertinente (Moerman et al., 2006). De la même manière, Van Sluis et al. décrivent dans leur revue de la littérature que les évaluations des voix après laryngectomie totale sont imparfaites et montrent un faible niveau de preuve. Il existe donc une urgence quant à la création d'outils de mesure standardisés pour l'évaluation des patients utilisant des voix de substitution (van Sluis et al., 2018).

1.3. Parole et rythme

Comme décrit précédemment, les paramètres objectifs acoustiques et aérodynamiques de la voix oro-œsophagienne ont été abondamment étudiés dans la littérature. Or, ces paramètres n'étant pas les plus adaptés pour évaluer la voix oro-

œsophagienne, il nous paraissait intéressant d'étudier la parole oro-œsophagienne, qui n'a fait l'objet que de peu d'études, et d'aborder plus précisément l'aspect rythmique des perturbations induites par la laryngectomie totale.

La parole est définie par le Centre National de Ressources Textuelles et Lexicales comme la *« faculté d'exprimer et de communiquer la pensée au moyen du système des sons du langage articulé émis par les organes phonateurs* ». Elle est une manifestation de la langue dans ses aspects acoustiques, physiologiques et cognitifs (*PAROLE : Définition de PAROLE*, s. d.).

Les perturbations de la parole englobent deux aspects : la production et la perception. Les organes phonateurs ont été fortement impactés et remaniés chez les patients ayant été opérés d'une laryngectomie totale. La perte radicale de la phonation laryngée engendre des conséquences sur leur parole, et cet ensemble de modifications handicape lourdement leur communication. Celle-ci présente toujours les mêmes contraintes qu'avant l'opération : s'exprimer de manière à transmettre un message qui sera compris par l'interlocuteur, et le sujet utilisant une voix oro-œsophagienne devra s'adapter afin de compenser ses troubles.

Un des aspects majeurs de la parole est son organisation temporelle – son rythme – qui influence l'intelligibilité du message au travers de la prosodie. Ce rythme implique différents paramètres comme l'accentuation, le timbre, le débit et les pauses. Les patients laryngectomisés totaux présentent des contraintes physiologiques qui perturbent le débit. En effet, afin de recharger en air le réservoir œsophagien, ils sont obligés d'effectuer des pauses de blocage, s'ajoutant aux pauses présentes dans la parole non-pathologique. La capacité limitée de ce réservoir et la réalisation effective du blocage vont donc soumettre la parole oro-œsophagienne à une réorganisation temporelle.

1.3.1. Rythme et pauses dans la parole normale

Le rythme est une composante fondamentale à toute activité motrice. La parole en représentant un, il lui est donc essentiel (Ordin, 2019). La prosodie, et plus particulièrement le rythme, contribuerait en effet à la compréhension du message par le locuteur au même titre que la parole en elle-même (Lalain, 2016 ; Lalain, 2012). De plus, la synchronisation rythmique de la parole entre deux locuteurs représenterait un facteur de réussite de l'interaction. Plus le degré de cette synchronisation serait élevé, meilleure serait la communication (Stephen. J. Cowley, 1994).

Le rythme de la parole est créé par l'alternance d'accélération et de ralentissement articulatoires, entrecoupée par l'ajout de pauses. Lorsque nous parlons, nous parsemons notre discours de ces interruptions plus ou moins longues (Ramanarayanan et al., 2009). Dans la littérature, on retrouve plusieurs catégorisations de ces pauses.

On distingue dans un premier temps les pauses pleines des pauses vides, ou silencieuses. Les pauses pleines représentent un « *arrêt non silencieux et momentané de l'avancement du discours* » (Hirsch et al., 2018), se traduisant par l'adjonction d'un « euh » avant ou après un mot. Elles marquent généralement une hésitation ou une planification de

l'énoncé (Duez, 2001). Lorsqu'elles sont silencieuses, elles se définissent comme une interruption de la phonation (Grosman et al., 2018 ; Hirsch et al., 2018). Ces pauses vides ne le seraient toutefois pas vraiment, car elles permettraient une activité respiratoire (Fuchs et al., 2008 in Hirsch et al., 2018) ou encore la préparation de mouvements articulatoires (Abry et al. 1996 in Hirsch et al., 2018).

Dans les différentes études, aucun consensus n'a pu être établi quant à la durée minimum des pauses. On commence en effet généralement à considérer une pause silencieuse dès lors que l'arrêt dans le flux de parole dépasse 200ms (Candea, 2000 ; Campione & Veronis, 2002 ; Campione & Veronis, 2005), mais ce seuil varie en fonction des auteurs : pour Grosman et al., il varie entre 180 et 300ms (Grosman et al., 2018) ; pour F. Grosjean et A. Deschamps à partir de 250ms (Grosjean et Deschamps, 1975). Dans le corpus étudié, seuls Lalain et al. émettent une limite supérieure, constatant qu'il est rare de retrouver des pauses de plus de 2 secondes dans la parole normale (Lalain et al., 2012).

Divers auteurs ont tout de même mis en avant le fait que la mise en place de seuils de durée pouvait modifier de manière très importante les résultats obtenus et donc leur interprétation. L'application de seuils fixes n'est donc pas préconisée pour exclure certaines données (Grosman et al., 2018). Campione et Véronis en 2002 ont comparé leurs résultats sans seuils limites et après mise en place de seuils inférieurs et/ou supérieurs. En ignorant les valeurs de durée extrêmes (très brèves ou très longues), ces seuils peuvent conduire à des conclusions totalement fausses : la durée des pauses est plus brève en parole spontanée qu'en parole lue sans aucun seuil, les durées moyennes sont relativement semblables entre les 2 tâches avec un seuil minimal de 200 ms et un seuil maximal de 2 000 ms, et les pauses deviennent plus brèves en parole lue qu'en parole spontanée si on n'applique que le seuil maximal de 2 000 ms (Campione & Véronis, 2002).

Se manifestant donc généralement par une rupture plus ou moins longue du signal acoustique, les pauses peuvent endosser différentes fonctions au sein du discours. Dans son étude, Candea parle de pause structurante ou non-structurante : l'une représente un élément, une étape dans le rythme prévu par le locuteur. L'autre fait comme irruption dans le discours, et est précédée d'une marque de travail de formulation (répétition, allongement d'hésitation, faux départ, auto-correction…) (Candea, 2000). Une autre terminologie est proposée par Ramanarayanan et al., avec des pauses dites grammaticales et non-grammaticales. Les premières seraient contrôlées cognitivement, faisant partie intégrante de la planification des mouvements articulatoires. Elles induiraient un ralentissement articulatoire à l'approche d'une frontière syntaxique, et seraient donc phonologiquement planifiées pour permettre au locuteur de varier le rythme de sa parole. A l'inverse, les pauses non-grammaticales apparaissent de manière inattendue, voire inappropriée, dans la phrase et viennent en perturber l'écoulement rythmique (Ramanarayanan et al., 2009).

Les pauses n'apparaissent donc pas de façon hasardeuse, et elles peuvent prendre différentes formes selon leur fonction dans le discours. Grosjean et Deschamps, en 1975, ont émis des réserves quant à leur classification, disant que tenter d'en déterminer précisément la nature ne semblait pas tâche aisée (Grosjean et Deschamps, 1975). Toutefois, en nous appuyant sur la littérature ainsi que sur la thèse de Camille Galant, nous avons sélectionné un panel de fonctions que peuvent revêtir les pauses au sein des différents registres de parole.

Certaines ont une visée *physiologique*, parmi lesquelles on retrouve des pauses respiratoires, permettant la reprise d'air, et des pauses servant à la déglutition (Dodane & Hirsch, 2018). Généralement, le locuteur tente de les placer dans son discours de manière à en perturber le moins possible le flux.

Les pauses *syntaxiques* s'apparentent aux pauses dites grammaticales citées précédemment, et s'utilisent volontairement à l'intérieur d'un tour de parole, pour structurer le discours. Elles induisent d'ailleurs généralement un ralentissement puis une accélération du débit de parole avant et après la frontière de la phrase, et pourrait être le lieu d'une respiration ou de la préparation du geste phonatoire suivant (Ramanarayanan et al., 2009). Elles coïncident généralement avec la reprise inspiratoire (Laroche-Bouvy, 1984).

Les pauses à fonction *cognitive*, enfin, sont majoritairement appelées pauses d'hésitation dans la littérature (Grosjean & Deschamps, 1975 ; Laroche-Bouvy, 1984 ; Candea, 2000 ; Hirsch et al., 2018). Elles servent à la recherche lexicale ou à la planification de l'énoncé (Hirsch et al.,2018), et peuvent être vides ou remplies. Généralement, le locuteur la comble avec un phonème (« euh »), afin de ne pas perdre son tour de parole (Laroche-Bouvy, 1984). La pause peut également servir à la focalisation, mettant en exergue une partie du discours (Hirsch et al., 2018). Puis la pause dites d'auto-correction a quant à elle un rôle en cas de mauvais départ ou d'accroc dans le discours (Nespoulous, 2014 in Dodane & Hirsch, 2018).

Toutefois, ces fonctions sont parfois difficiles à déterminer lors de l'analyse des pauses du discours. En effet, on note régulièrement la co-occurrence de différentes fonctions au sein d'une même pause, si bien qu'une fin de phrase est l'occasion d'une respiration, ou une respiration le prétexte d'une hésitation (Grosjean & Deschamps, 1975).

En termes de durée, il semble que les pauses soient plus longues lors de tâches telles que la description d'image que lors de la lecture (Hirsch et al., 2018). Les pauses silencieuses peuvent être divisées en trois classes : courtes (lorsqu'elles sont inférieures à 200 ms), moyennes (entre 200 et 1000 ms) et enfin longues (avec une durée supérieure à 1000 ms). Dans une étude de Campione et Véronis, comparant la durée des pauses en lecture et en spontané dans différentes langues, il est ressorti que les pauses dites longues n'étaient observées qu'en spontané (Campione & Véronis, 2002). Il semble donc que le degré de préparation de la parole avant une production soit un facteur influençant la durée des pauses qui s'adjoindront au discours (Grosman et al., 2018). L'endroit où se fait la pause pourrait également en impacter la durée : on note un temps de pause moyen plus long en fin d'énoncé (881 ms) ou de proposition (608 ms) que lorsqu'elle se fait au milieu d'une phrase (486 ms) (Grosman et al., 2018). D'ailleurs, 94% des pauses dites longues se trouvent en position finale (Goldman, 2010).

Les différences de durée observées pourraient enfin également être dues à la présence de différentes fonctions au sein d'une même pause. On a vu en effet qu'une même interruption de parole pouvait être utilisée pour assurer plusieurs besoins en simultané (Grosjean & Deschamps, 1975). Il semble alors que l'adjonction de plusieurs fonctions au sein de la pause en influence sa durée : une pause syntaxique seule est plus courte que lorsqu'elle s'associe à une fonction physiologique (prise de souffle, déglutition, etc…) (Goldman, 2010 ;

Hirsch et al., 2018). En termes de chiffres, une pause syntaxique seule aurait en moyenne une durée de 343 ms, une pause non-syntaxique (type hésitation) 617 ms. Lorsqu'elles sont couplées avec une fonction physiologique, les durées peuvent monter jusqu'à plus de 1000 ms (Hirsch et al., 2018).

1.3.2. Perturbation du rythme de la parole

Les pauses sont donc des entités présentes naturellement dans notre discours. Toutefois, le rythme de notre parole peut être altéré dans diverses pathologies. Le bégaiement en est un exemple bien connu : il s'agit d'un trouble de la fluence du discours, dans lequel on trouve un pourcentage de pauses silencieuses plus important que dans la parole normale (Teixeira, 2012). Dans la dyslexie, on retrouve des pauses plus longues que chez les sujets sains. Cette différence pourrait être expliquée par le coût cognitif plus important qu'engendre la parole chez les patients dyslexiques. Dans la littérature, ils sont d'ailleurs décrits comme étant les seuls à produire des pauses « intra-mot », dues à des difficultés de décodage (Lalain, 2012). Borrie, dans son étude de 2018, a montré que le rythme de la parole était atteint dans la dysarthrie (Borrie et al., 2018). Les maladies neurodégénératives induisent également des modifications de la parole. Dans la maladie de Parkinson, on observe une augmentation de l'occurrence de pauses, consécutive à une atteinte à la fois motrice et cognitive (Smith & Caplan, 2018). Dans la sclérose latérale amyotrophique, Yusunova et al. ont mis en avant des durées de pauses et de production de phrases allongées, surtout dans le cas de troubles respiratoires. Cela induirait également une diversité moins importante des types de pauses (L. Lee et al., 1993 in Yusunova et al., 2016). Enfin, la maladie d'Alzheimer engendre également un changement rythmique de la parole. A la fois dû à des difficultés d'accès lexical et un ralentissement cognitif (Nef & Huet, 1992 in Barkat-Defradas, 2009), cela implique une augmentation du nombre de pauses ainsi que de leur durée dans le discours (Barkat-Defradas, 2009).

La production du rythme de la parole dans ces différentes pathologies est donc entravée, malgré la présence de tous les organes nécessaires à la phonation. Or, son altération peut également découler d'un trouble mécanique. Après laryngectomie totale, le sujet est privé de son larynx, et donc de ses cordes vocales. Le patient va donc apprendre à produire la parole autrement (Graville et al., 2018). Il existe différents moyens de phonation, dont l'utilisation de la voix oro-œsophagienne, qui induit l'adjonction de nombreuses pauses. Celles-ci constituent une nécessité physiologique, car pour produire du son, le sujet doit bloquer de l'air dans le réservoir œsophagien (Zenga et al., 2018). Si l'on a pu voir précédemment que la voix oro-œsophagienne avait fait l'objet de diverses études, la parole oro-œsophagienne n'a toutefois été que peu disséquée.

1.3.3. Rythme et pauses dans la parole oro-œsophagienne

Globlek D. et al., ont analysé la qualité de prononciation des différents utilisateurs de voix alaryngées, et les patients utilisant la voix oro-œsophagienne se sont révélés les moins bons (Globlek et al., 2004). En effet, leur parole était jonchée de disfluences et de bruits sourds

consécutifs aux altérations des paramètres acoustiques de leur voix. D'après Bellandese, les orateurs oro-œsophagien produisent plus de pauses que les sains (22 pour 8.2) (Bellandese et al., 2001). On le sait, pour générer une phonation, le sujet laryngectomisé procède à un blocage d'air dans l'œsophage, générant donc une pause dans le discours. En moyenne, il peut produire sept syllabes par prise d'air – alors que le sujet sain en produira 19 – ce qui explique la multiplication des pauses dans leur parole (van Rossum, 2008). Or, la quantité de pauses, influençant la fluidité du discours (Moerman et al., 2015), représenterait un facteur d'intelligibilité de la parole oro-œsophagienne, et les bons parleurs en produiraient moins (Motta et al., 2001). De plus, leur caractère physiologique les rend inévitables, voire incontrôlables, car elles peuvent être nécessaires à n'importe quel moment de la phonation. De fait, là où le sujet sain tente de placer ses pauses physiologiques de manière appropriées, le parleur oro-œsophagien n'en a pas la possibilité. On remarque également un effet de longueur : plus la phrase est longue, plus on retrouve de pauses inappropriées (van Rossum, 2008).

Les pauses sont donc au cœur des altérations rythmiques retrouvées dans la parole oro-œsophagienne. Or, la plupart des études conclut effectivement à une plus grande quantité de pauses que chez le sujet sain, sans toutefois en avoir étudié plus spécifiquement la nature et la disposition dans la parole des patients laryngectomisés. Après un état des lieux de la question, Van Rossum est l'une des seuls à avoir approfondi les recherches (van Rossum et al., 2002, 2005, 2008). Dans son étude, elle a effectivement constaté l'adjonction de nombreuses pauses venant perturber le rythme du discours. Mais celles-ci étaient généralement localisées au niveau de frontières prosodiques, remplaçant l'allongement de la syllabe finale normalement retrouvé chez le sujet sain. Cette observation est d'ailleurs retrouvée dans une analyse des marqueurs de frontières prosodiques faite chez des locuteurs laryngectomisés totaux hispanophones, où la pause semble également se substituer à l'allongement prosodique final (Cuenca & Barrio, 2011). Van Rossum a également remarqué un effet de longueur dans la parole oro-œsophagienne : en effet, il semble que plus la phrase est longue, plus il y a de pauses inappropriées. Toutefois une utilisation plus cohérente des pauses est observée plus le niveau de complexité augmente. Enfin, la présence de consonnes plosives (k, p, ou encore b) peuvent permettre une injection d'air discrète au milieu d'une phrase, et permettrait donc au locuteur alaryngé d'économiser une pause intra-phrase (van Rossum, 2008). Cette étude ne se rapporte néanmoins qu'à la langue néerlandaise, et ne permet pas d'établir une généralisation de ces observations.

Plus récemment, Camille Galant a offert dans son travail de thèse une analyse des pauses dans le cadre d'une tâche de lecture, en langue française, comparant sujets laryngectomisés et sujets sains. Les résultats ont corroboré les précédentes études quant à la quantité de pauses plus élevée chez les utilisateurs d'une voix oro-œsophagienne. En effet, la sonorisation de leur parole n'est possible qu'au détriment de la fluidité de cette dernière. Elle retrouve autant de pauses syntaxiques chez les laryngectomisés que chez les sains, ce qui semble s'expliquer par l'utilisation des frontières syntaxiques du texte pour réaliser leurs blocages. Toutefois, les pauses de blocage sont in fine les plus représentées dans la collecte des données. En comparant les durées des différentes pauses, celles à fonction syntaxique sont plus longues que les blocages simples. Elle met en avant deux profils chez les sujets alaryngés : ceux dont la durée des pauses de blocages est égale à celle des syllabes et ceux dont les blocages sont plus longs que les syllabes. Dans le deuxième cas, on pourrait imaginer

que cette différence de durée s'explique par un désir de remplir au maximum les réservoirs d'air et donc réduire le nombre de pauses. Mais aucune différence significative n'a été retrouvée quant à la quantité des pauses chez les deux profils. Enfin, contrairement à Van Rossum et à Cuenca et Barrio, on retrouve un allongement de la syllabe finale aux frontières prosodiques chez les sujets laryngectomisés (Galant, 2019). Cela peut être mis en lien avec la langue étudiée. En effet, l'allongement final est caractéristique du français et semblerait persistant même dans une atteinte de la parole.

Aux différents types de pauses évoqués précédemment, on peut alors ajouter pour analyser la parole oro-œsophagienne la pause physiologique de blocage, mais aussi la pause d'injection. Celle-ci consiste à utiliser une consonne (plosive ou occlusive) pour injecter de l'air (van Rossum, 2008).

Notre mémoire fait suite à la thèse de Camille Galant, et s'intéresse donc aux altérations rythmiques consécutives à la laryngectomie totale. Pour notre analyse, nous étudierons la parole semi-spontanée, sur une tâche de description d'image. Nous savons que l'apparition des différents types de pauses est influencée par le degré de préparation du discours : préparé, semi-spontané ou spontané (Grosman et al., 2018). On constate par exemple plus de pauses à fonction cognitive dans une tâche de parole spontanée qu'en lecture (Lalain et al., 2016). En description d'image justement, les pauses, qu'elles soient silencieuses ou pleines, occuperaient 40 à 50% du temps de parole (Goldman-Eisler, 1968 in Hirsch et al., 2018).

Notre recherche permettrait de fournir une description phonétique de l'organisation temporelle de la parole oro-œsophagienne. En effet, notre mémoire portant sur l'étude du rythme de la parole du sujet oro-œsophagien par rapport à celle du sujet sain, nous avons comme objectif de recenser et catégoriser les pauses dans le discours de ces deux populations afin d'en apprécier la distribution et donc de mettre en évidence les modifications rythmiques induites par la laryngectomie totale. Au vu de l'adjonction de nombreuses pauses, nécessaires à la phonation, nous supposons que l'utilisation d'une voix oro-œsophagienne vient perturber la structuration temporelle du discours, et en altérer l'intelligibilité. De plus, nous voulons vérifier deux hypothèses : la première, confirmant les études précédentes, est que la quantité de pauses dans la parole du sujet laryngectomisé est plus importante que chez le sujet sain ; la seconde étant que ces pauses plus nombreuses varient en durée selon leur fonction.

2. Méthodologie de la recherche

Dans ce travail de mémoire, nous souhaitons réaliser une étude du type, de la distribution et de la durée des différentes pauses de la parole oro-œsophagienne afin de les comparer à celles de la parole des sujets sains.

2.1. Expérimentation

2.1.1. Population

La population étudiée comprenait :

- 20 sujets opérés d'une laryngectomie totale (moyenne d'âge 67,45 ans, médiane : 71,5 ans, Ecart-type : 10,5 ans, minimum : 42 ans, maximum : 83 ans). Ces sujets ont été inclus au cours d'une consultation de suivi fonctionnel au sein du service ORL des hôpitaux La Timone et La Conception entre mai 2015 et mai 2016. Cet échantillon comprenait 18 hommes (soit 90% de la population) et 2 femmes (soit 10% de la population). La distance moyenne à la chirurgie au moment des enregistrements était de 61, 8 mois. Le maximum étant de 252 mois et le minimum de 6 mois.
- 21 sujets sains appariés en sexe et en âge ont été enregistrés, en comparaison des sujets pathologiques (moyenne d'âge 67 ans, médiane : 66 ans, Ecart-type : 7,1 ans, minimum : 57 ans, maximum : 82 ans). L'ensemble des sujets était domicilié en région PACA.

Le tableau récapitulatif des sujets pathologiques et sains est disponible en annexe 1 et 2.

Critères d'inclusion :

- Avoir bénéficié d'un traitement par laryngectomie ou pharyngo-laryngectomie totale.
- Avoir choisi la voix oro-œsophagienne comme mode de phonation.
- Avoir une maîtrise de la voix oro-œsophagienne suffisante pour communiquer dans une situation duelle, sans l'aide d'un tiers.

Critères de non-inclusion :

- Sujet présentant un trouble neurologique associé.
- Sujet présentant ou ayant présenté un trouble de la parole ou du langage.
- Sujet présentant un trouble de l'audition.
- Sujet ayant subi une modification des organes bucco-phonatoire autres que l'exérèse du larynx.
- Sujet ne maîtrisant pas la langue française.

2.1.2. Corpus

Le corpus utilisé dans ce travail de mémoire comprenait des échantillons de voix oro-œsophagienne et saine en modalité semi-spontanée, dans le cadre d'une tâche de description d'image.

2.1.2.1. Recueil de parole

La banque d'images utilisée pour cette tâche de description a été réalisée dans le cadre d'un mémoire d'orthophonie étudiant le discours des sujets ayant subi une laryngectomie totale, dirigé en 2016 par Muriel Lalain et Camille Galant. La parole semi-dirigée implique des contraintes lexicales tout en permettant à la personne de s'exprimer de manière plus libre qu'en tâche de lecture. 10 photos constituent la banque d'images, sur lesquelles sont représentés des personnages LEGO® dans un décor. Ces photos se distinguent les unes des autres par différents paramètres : l'objet que les personnages tiennent à la main (une lettre, une clé, un balai ou une mallette) et la couleur d'un livre présent en arrière-plan. Deux sets de ces 10 images ont été produits.

Le sujet avait pour consigne de décrire une carte, piochée au hasard dans son jeu, en donnant le plus de détails possibles pour que l'examinateur retrouve le double dans son propre jeu. L'examinateur donnait alors pour consigne : *« choisissez une carte dans votre jeu et décrivez là le plus précisément possible afin que l'examinateur puisse la retrouver au sein de son propre jeu »*

Un exemple d'images est disponible en annexe 4.

2.2. Traitement des données

Le traitement des données effectué concernait les caractéristiques temporelles de la parole en voix oro-œsophagienne, obtenues à partir des échantillons de discours semi-dirigé recueillis dans le corpus créé dans la thèse de Camille Galant.

2.2.1. Transcription orthographique

Pour étudier leur parole, nous avons dans un premier temps procédé à une transcription orthographique exacte des productions des différents sujets. Dans le but d'éviter des erreurs de transcription, la méthodologie utilisée comprenait 3 expérimentateurs et plusieurs étapes :

1 – Deux expérimentateurs A et B, qui étaient deux étudiants en orthophonie, ont écouté et transcrit individuellement chaque enregistrement

2 – Un troisième expérimentateur C, orthophoniste spécialisée dans la prise en charge des patients opérés d'une laryngectomie totale, a alors vérifié les transcriptions des deux expérimentateurs A et B et a souligné les différences de transcription

3 – Les expérimentateurs A et B ont alors écouté à nouveau les enregistrements en portant une attention particulière aux passages discordants. Ils ont ensuite proposé une nouvelle transcription ou maintenu leur transcription initiale.

4 – L'expérimentateur C a décidé de la bonne transcription s'il persistait des désaccords. C'est cet expérimentateur qui a alors défini la transcription définitive.

Les transcriptions orthographiques ont été réalisées pour les tâches de descriptions d'image.

2.2.2. Segmentations

Le travail de thèse de Camille Galant ayant uniquement porté sur la lecture, nous avons poursuivi cette expérimentation en étudiant dans notre travail de mémoire le discours semi-spontané.

Une segmentation de la parole étudiée a été nécessaire, consistant à placer des frontières entre certains segments de parole. Nous avons procédé à une segmentation semi-automatique des enregistrements à partir des transcriptions faites. Cette dernière a été réalisée avec le logiciel libre Praat© (Boersma & Weenink, 2001) et le plug-in Easy Align© (Goldman, 2011). Le plug-in a réalisé une pré-segmentation à 5 niveaux (appelés tier) des enregistrements à partir de la transcription orthographique qui lui a été fournie.

Cette pré-segmentation a ensuite été corrigée manuellement par les expérimentateurs A et B, qui ont également annoté chaque pause du discours des sujets du corpus après discussion et vérifications avec l'expérimentateur C. Les niveaux de segmentation étaient les suivant :

- La tier 1 était la tier orthographique. A ce niveau l'échantillon de parole a été segmenté selon les phrases composant le discours semi-dirigé.
- La tier 2 était la tier phonologique. A ce niveau, l'échantillon de parole a été également segmenté selon les phrases du discours, mais il a été automatiquement retranscrit en utilisant la convention SAMPA. Cette convention permet une transcription phonétique utile pour la suite des analyses.
- La tier 3 était la tier mots. Le plug-in Easy Align© a alors segmenté chaque échantillon de parole en mots. Il a donc placé de manière automatique des frontières entre chaque mot produit par le sujet. Chaque interruption de parole, caractérisée par une absence de signal de parole a été automatiquement étiquetée « _ » par le logiciel.
- La tier 4 était la tier syllabes. Le plug-in Easy-Align a alors segmenté le signal de parole par syllabe.
- La tier 5 était la tier phonèmes. Le plug-in Easy-Align a alors segmenté le signal de parole phonème par phonème.

Les expérimentateurs ont révisé et corrigé manuellement la segmentation semi-automatique en s'aidant des indices du spectre de la parole pour positionner les frontières entre les éléments selon la tier. La segmentation a respecté les principes de segmentation décrits par Machah P, et Skarnitzl R. dans leur ouvrage Principle of phonetics segmentations (Skarnitzl & Machač, 2009). En cas de difficulté, l'expérimentateur C est intervenu. Cette procédure a eu lieu individuellement pour chaque sujet LT et sain.

La tier 4, tier syllabes, a été dupliquée manuellement pour chaque sujet afin d'annoter les syllabes précédant les pauses syntaxiques étiquetées _psy. Ces syllabes ont été étiquetées _syllfinale. Elles nous permettront d'évaluer un possible allongement de la dernière syllabe.

La méthodologie de correction de la pré-segmentation et de l'annotation des pauses était la suivante :

- L'expérimentateur A a corrigé et annoté la moitié des sujets et l'expérimentateur B l'autre moitié.
- Ils ont ensuite échangé leurs sujets afin de procéder à une première vérification de leurs productions.
- La segmentation et l'annotation définitives pour chaque sujet LT et chaque sujet sain ont été validées après discussion et vérifications au cas-par-cas avec l'ensemble des expérimentateurs (A, B et C).

2.2.3. Etiquetage des pauses

L'objectif de l'expérience était d'analyser les pauses présentes dans la parole des sujets opérés d'une laryngectomie ou pharyngo-laryngectomie totale. A la suite de la phase de segmentation, nous avons annoté chaque pause trouvée dans la parole de nos sujets. Comme nous avons pu le stipuler précédemment, nous avons mené cette expérimentation sur de la parole semi-spontanée. Ce type de parole présentant moins de contraintes syntaxiques que la lecture, et comprenant des étapes de programmation sémantique, nous nous attendions à une plus grande variété de pauses.

Comme vu précédemment dans notre état des lieux, plusieurs auteurs ont mis en évidence que l'ajout de seuils de durée aurait pour conséquence de modifier de manière importante les résultats obtenus ainsi que leur interprétation. Dans notre étude nous avons fait le choix de ne pas comptabiliser les pauses inférieures à 200 ms, sauf dans les cas où une fonction était audible (prespi, phesr, pbloc, etc). Dans ce cas, toutes les pauses même inférieures à 200 ms étaient comptabilisées. En effet, bien que brèves, ces pauses physiologiques n'en étaient pas moins perturbatrices de la parole. Aucun seuil supérieur n'a été appliqué, seules quelques valeurs extrêmes (9 pauses de blocages syntaxiques supérieures à 3 secondes) ont été supprimées des analyses statistiques.

Nous avons divisé les différents types de pauses en 3 grandes catégories : les pauses physiologiques, les syntaxiques et les cognitives. Dans chacune d'elles, nous retrouvions toutes sortes de pauses : respiratoires, syntaxiques, d'hésitation, de focalisation, de blocage, etc. Nous considérions comme pause d'hésitation toute interruption du discours servant à la planification et l'élaboration du discours, générant une hésitation dans le flux de parole. Or, il semblait que ces différents types pouvaient s'entremêler, devenir le prétexte d'une autre fonction : une pause syntaxique pouvait être l'occasion d'une respiration ou d'un blocage, un blocage pouvait s'ajouter à une hésitation, etc… (Grosjean & Deschamps, 1975). Et ce aussi bien chez le sujet sain que chez le sujet laryngectomisé. De fait, nous avons créé des noms pour spécifier ces pauses à double fonction (pblocsy, prespisy, pbloches, etc…). Nous avons également nommé "tour de parole" (_tdp) les interruptions permettant le changement de locuteur.

Les sujets ont réalisé diverses pauses, pour lesquelles nous avons réalisé un étiquetage spécifique. Ces pauses ont été d'abord détectées automatiquement par le logiciel Praat©, et

annotées à l'aide du signe « _ ». Puis, chacune a été nommée selon sa fonction au sein de la parole.

Les différents types de pauses sont répertoriés dans le tableau ci-dessous.

Tableau 3.1 : Les différents types de pauses étiquetées dans le corpus

Etiquetage	Type de pause
_psy	Pause syntaxique
_pbloc	Pause de blocage (sujet LT)
_pblocsy	Pause de blocage à fonction syntaxique (sujet LT)
_pac	Pause d'autocorrection
_pblocac	Pause d'autocorrection avec blocage (sujet LT)
_psf	Pause de focalisation
_prespi	Pause de reprise inspiratoire (sujet LT ou sain)
_prespisy	Pause syntaxique donnant lieu à une reprise inspiratoire (sujet LT ou sain)
_pdeglu	Pause de déglutition (sujet LT ou sain)
_pinj	Pause où le sujet LT injecte de l'air via l'ajout d'une consonne (occlusive ou plosive) sans blocage audible
_pinjsy	Pause syntaxique donnant lieu à une injection (sujet LT)
_pinjhes	Pause d'hésitation avec injection (sujet LT)
_phes	Pause d'hésitation vide (sujet LT ou sain)
_phesr	Pause d'hésitation remplie, généralement par un "euh" (sujet LT ou sain)
_pbloches	Pause d'hésitation avec blocage (sujet LT)
_prespihes	Pause d'hésitation permettant une reprise inspiratoire (sujet LT ou sain)
_bbouche	Bruit de bouche ne correspondant ni à une déglutition, ni à un blocage (sujet LT ou sain)
_pi	Pauses idiopathiques (non catégorisées)
_tdp	Tour de parole, présentes entre les tours de parole du sujet et de l'orthophoniste réalisant l'enregistrement

Les blocages entendus dans un tour de parole n'étaient pas pris en compte, car ils ne représentaient pas réellement une pause au sein de l'énoncé. Nous avons pu observer chez nos sujets laryngectomisés l'occurrence successive de pauses de respiration et de blocage. Nous avons décidé de les rassembler et de les nommer _pbloc, car nous estimions que l'inspiration entendue faisait partie du blocage. De fait, une pause de blocage syntaxique *précédée* d'une pause respiratoire était annotée "_pblocsy". Chez ces sujets, on retrouvait d'ailleurs très souvent de longues pauses, avec différentes fonctions présentes en une. Or, déterminer précisément où commence et s'achève chaque type de pauses au sein même d'une interruption était impossible, surtout lorsque l'on prenait en compte le fait que certaines peuvent servir à la réalisation d'une autre.

Nous voulions prouver que les pauses de blocages simples, à visée purement physiologique, étaient généralement plus courtes que lorsqu'elles étaient concomitantes à une autre fonction, d'où notre décision de les rassembler. Dans ces pauses associant plusieurs fonctions (pblocsy, pbloches), le blocage ne s'entendait en effet pas sur toute la pause, mais

son caractère physiologique le rendait obligatoire à la reprise de la phonation. Pour les pauses syntaxiques, nous les positionnions dès qu'une pause semblait correspondre à un signe de ponctuation (virgule, point, etc…) et que cela semblait structurer le discours.

Enfin, pour différencier une pause d'auto-correction d'une pause d'hésitation, nous avons décidé la chose suivante : la première était choisie en cas de changement de mot ("un… deux personnages") ou de rectification de la prononciation, tandis que la seconde coupait le discours mais ne changeait pas l'énoncé en cours.

2.2.4. Extraction des données

L'extraction des données a été réalisée via l'utilisation d'un script Praat© créé par Céline De Looze, mis à disposition de la communauté scientifique sur son site internet (http://celinedelooze.com/lien_outils.htm). Les données ont été extraites à partir de la tier syllabe dupliquée (contenant les « _syllfinale ») et de la tier phonème de nos segmentations. Ces données comprenaient les durées de segments : syllabes, phonèmes et pauses annotées.

2.3. Analyses statistiques

Les analyses statistiques ont été réalisées via le logiciel R (R Core Team, 2013). (R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/.v Version : 4.0.5).

Notre échantillon de sujets opérés d'une laryngectomie totale étant de petite taille, et n'ayant pas une distribution normale, nous avons réalisé des tests non paramétriques, notamment le test de Wilcoxon. Nos résultats ont été considérés comme significatifs à partir d'une p-valeur inférieure ou égale à 0,05 signifiant un intervalle de confiance à 95%. Les p-valeurs exactes seront mentionnées.

3. Résultats

3.1. Analyse des pauses

3.1.1. Quantité

Dans un premier temps, nous avons extrait les différents segments de parole (pauses, syllabes ou tour de parole) que l'on retrouve pour les deux groupes de sujets. Nous avons répertorié leur occurrence dans le tableau ci-dessous.

Tableau 3.1 : Occurrences des différents segments de parole

Type de segment	Occurrence
Syllabes	6165
Pauses de blocages simples (pbloc)	489
Pauses de blocages à fonction syntaxique (pblocsy)	193
Pauses syntaxiques (psy)	50
Pauses respiratoires (prespi)	46
Pauses respiratoires avec fonction cognitive (prespihes)	45
Pauses respiratoires avec fonction syntaxique (prespisy)	175
Pauses d'hésitation vides (phes)	56
Pauses d'hésitation remplies (phesr)	190
Pauses d'hésitation avec blocage (pbloches)	45
Pauses d'autocorrection (pac)	22
Pauses d'autocorrection avec blocage (pblocac)	20
Pauses d'injection (pinj)	29
Pauses d'injection avec fonction syntaxique (pinjsy)	5
Pauses d'injection avec fonction cognitive (pinjhes)	1
Pauses de focalisation (psf)	6
Pauses de déglutition (pdeglu)	4
Pauses idiopathiques (pi)	2
Bruits de bouche (bbouche)	8

Comparaison de la quantité globale de pauses chez le sujet LT vs le sujet CTRL

Toutes pauses confondues, les sujets ayant été opérés d'une laryngectomie totale (LT) font en moyenne 43 pauses, tandis que les sujets sains n'en effectuent que 24 en moyenne.

Tableau 3.2 : Quantité moyenne de pauses chez les sujets LT et CTRL

	Nombre de sujets	Nombre de pauses	Ecart-type
LT	20	43,30	18,80
CTRL	21	24,43	15,41

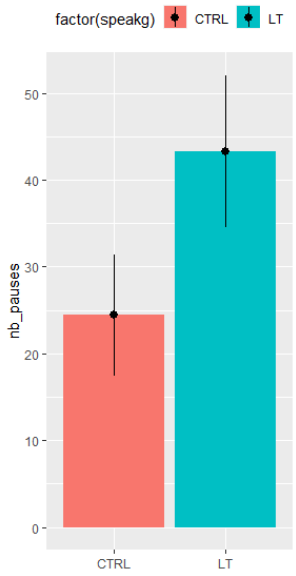

Figure 3.1 : quantité de pauses chez les sujets CTRL (rouge) et chez les sujets LT (bleu).

Après avoir effectué un test de Wilcoxon, nous trouvons une différence significative (p = 0,0007261). Les sujets LT font en moyenne plus de pauses que les sujets CTRL.

Comparaison de la quantité de pauses à fonction syntaxique chez le sujet LT vs le sujet CTRL

Nous avons comparé la quantité de pauses à fonction syntaxique (pblocsy, psy, prespisy, pinjsy confondues) retrouvée chez les sujets laryngectomisés par rapport à celle des sujets sains.

Tableau 3.3 : Quantité moyenne de pauses syntaxiques chez les sujets LT et CTRL

	Nombre de sujets	Nombre de pauses	Ecart-type
LT	20	11	6,27
CTRL	21	9,67	5,95

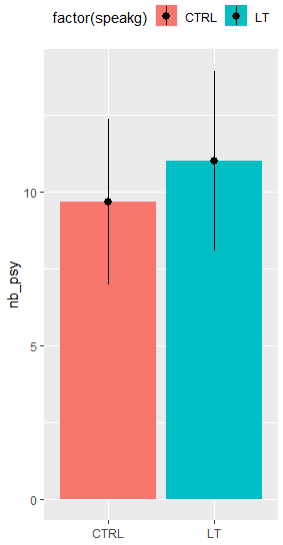

Figure 3.2 : Quantité de pauses syntaxiques chez les sujets CTRL (rouge) et chez les sujets LT (bleu)

La quantité moyenne de pauses à fonction syntaxique est équivalente chez les sujets laryngectomisés et chez les sujets sains. Le test de Wilcoxon effectué ne retrouve pas de différence significative entre les deux groupes de sujets (p = 0.454).

Comparaison de la quantité de pause d'hésitation chez le sujet LT vs le sujet CTRL

Nous avons comparé la quantité de pauses à fonction d'hésitation (phes, phesr, pbloches, prespihes et pinjhes confondues) retrouvée chez les sujets laryngectomisés par rapport à celle des sujets sains.

Tableau 3.4 : Quantité moyenne de pauses à fonction d'hésitation chez les sujets LT et CTRL

	Nombre de sujets	Nombre de pauses	Ecart-type
LT	20	4,70	3,20
CTRL	21	11,57	7,90

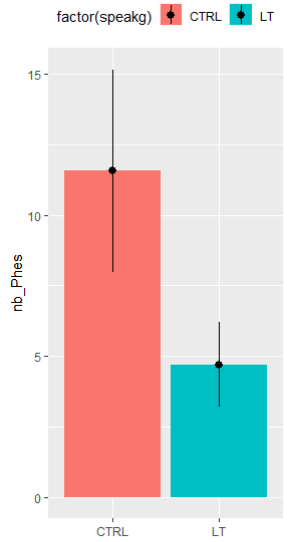

Figure 3.3 : Quantité de pauses d'hésitation chez les sujets CTRL (rouge) et chez les sujets LT (bleu)

On note que les sujets CTRL font en moyenne plus de pauses d'hésitation que les sujets LT, avec une moyenne de 12 pour les premiers et de 5 pour les seconds. Le test de Wilcoxon confirme que cette différence est significative (p = 0,0002765).

Comparaison de la quantité de pauses d'hésitation remplies chez le sujet LT vs le sujet CTRL

Ici, nous comparons la quantité de pauses d'hésitation remplies (phesr) réalisées par les sujets sains par rapport aux sujets opérés.

Tableau 3.5 : Quantité moyenne de pauses d'hésitation remplies chez les sujets LT et CTRL

	Nombre de sujets	Nombre de pauses	Ecart-type
LT	20	1,45	1,70
CTRL	21	7,67	6,94

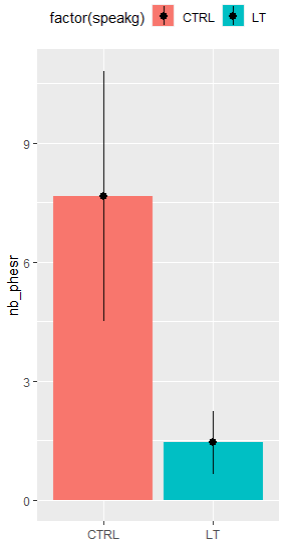

Figure 3.4 : Quantité de pauses d'hésitation remplies chez les sujets CTRL (rouge) et chez les sujets LT (bleu)

Le test de Wilcoxon effectué montre une différence significative entre les deux populations (p = 2,732 e-0,5). Les sujets CTRL font plus de pauses d'hésitation remplies – en moyenne 8 – que les sujets LT – 1 en moyenne.

Comparaison de la quantité de pauses d'autocorrection chez le sujet LT vs le sujet CTRL

Nous avons observé l'occurrence de pauses d'autocorrection chez les sujets utilisant la voix oro-œsophagienne comparée aux sujets sains. Nous avons pris en compte les pauses d'autocorrection simples (pac) et les pauses d'autocorrection avec blocage (pblocac) dans ces statistiques.

Tableau 3.6 : Quantité moyenne de pauses d'autocorrection chez les sujets LT et CTRL

	Nombre de sujets	Nombre de pauses	Ecart-type
LT	20	1,25	1,21
CTRL	21	0,81	1,17

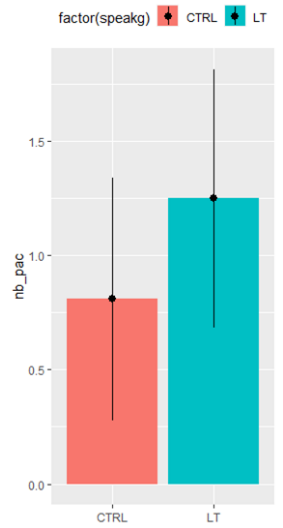

Figure 3.5 : Quantité de pauses d'autocorrection chez les CTRL (rouge) et chez les LT (bleu)

La quantité moyenne de pauses d'autocorrection retrouvée chez les sujets CTRL est équivalente à celle des sujets LT. Le test de Wilcoxon réalisé ne retrouve pas de différence significative entre les deux groupes (p = 0,175).

Comparaison de la quantité de pauses de blocage vs pauses d'injection chez les sujets LT

Nous nous sommes ici concentrées sur la différence de quantité entre les pauses de blocage simples (pbloc) et les pauses d'injection simples (pinj) que l'on retrouve chez le sujet LT. Nous en avons comparé l'occurrence dans le tableau ci-après.

Tableau 3.7 : Quantité moyenne de « pbloc » versus « pinj » chez les sujets LT

	Occurrence	Quantité moyenne/sujet	Ecart-type
pbloc	489	24,45	14,46
pinj	29	1,45	3,83

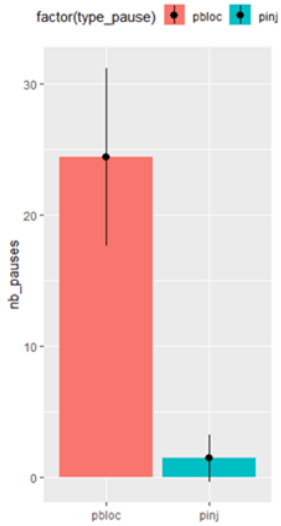

Figure 3.6 : Sujets LT : quantité des pbloc (rouge) et des pinj (bleu)

La différence est ici très significative, comme en témoigne le test de Wilcoxon (p = 0,0001505). Cela se retrouve aussi bien en termes d'occurrence – avec 489 pbloc relevées au sein des sujets LT contre 29 pinj – que de quantité moyenne par sujet. Les sujets LT font en moyenne 24 pauses de blocage lors d'une tâche de description d'image, contre 1 pause d'injection.

Comparaison de la quantité de pauses de blocage vs pauses de blocage à fonction syntaxique chez les sujets LT

Nous avons étudié la différence de quantité entre les pauses de blocage simples (pbloc) et les pauses de blocage syntaxiques (pblocsy) que l'on retrouve chez le sujet LT. Nous en avons comparé l'occurrence dans le tableau ci-après.

Tableau 3.8 : Quantité moyenne de « pbloc » versus « pblocsy » chez les sujets LT

	Occurrence	Quantité moyenne/sujet	Ecart-type
pbloc	489	24,45	14,46
pblocsy	193	9,65	4,91

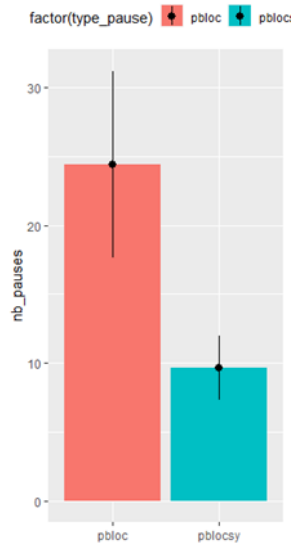

Figure 3.7 : Sujets LT : quantité de pbloc (rouge) et de pblocsy (bleu)

Les sujets LT produisent plus de pauses de blocage simple que de pauses de blocage à fonction syntaxique. Ils font en moyenne 24,45 pauses de blocage lors d'une tâche de description d'image, contre 9,65 pauses de blocage syntaxiques. La différence est significative, comme en témoigne le test de Wilcoxon (p = 0,0004786).

3.1.2. Durée

Comparaison de la durée des pauses d'injection simples vs pauses de blocage simples.

Nous avons comparé la durée entre les pauses d'injection (pinj) et les pauses de blocage (pbloc) chez les sujets laryngectomisés.

Tableau 3.9 : Durée moyenne de « pinj » versus « pbloc » chez les sujets LT

	Occurrence	Durée	Ecart-type
pbloc	489	0,48	0,33
pinj	29	0,36	0,16

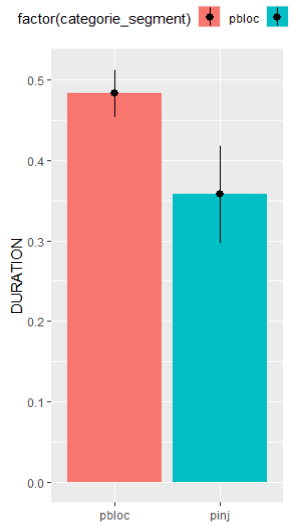

Figure 3.8 : Sujets LT : durée des pbloc (rouge) et des pinj (bleu)

La durée des pauses d'injection simples est équivalente à la durée des pauses de blocage simples. Le test de Wilcoxon réalisé ne retrouve pas de différence significative entre les deux types de pauses (p = 0,141).

Comparaison de la durée des pauses d'injection à fonction syntaxique vs pauses de blocage à fonction syntaxique.

Nous avons réalisé la comparaison entre la durée des pauses d'injection syntaxiques (pinjsy) et celle des pauses de blocage syntaxiques (pblocsy). Neuf valeurs extrêmes ont été supprimées des analyses statistiques, il s'agit de pauses de blocage à fonction syntaxique dont la durée était supérieure à trois secondes.

Tableau 3.10 : Durée moyenne de « pinjsy » versus « pblocsy » chez les sujets LT

	Occurrence	Durée	Ecart-type
pblocsy	184	1,09	0,59
pinjsy	5	0,67	0,17

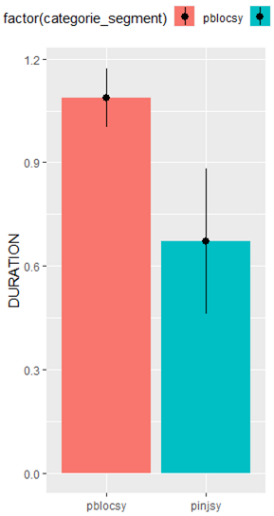

Figure 3.9 : Sujets LT : durée des pblocsy (rouge) et des pinjsy (bleu)

La durée des pauses d'injection syntaxiques est équivalente à celle des pauses de blocage syntaxiques. Le test de Wilcoxon réalisé ne retrouve pas de différence significative entre les deux types de pauses (p = 0,09098).

Comparaison de la durée des pauses de blocage simples vs pauses de blocage à fonction syntaxique.

Nous avons réalisé la comparaison entre la durée des pauses de blocage simples (pbloc) et celle des pauses de blocages syntaxiques (pblocsy). 9 valeurs extrêmes ont été supprimées des analyses statistiques, il s'agit de pauses de blocage à fonction syntaxique dont la durée était supérieure à 3 secondes.

Tableau 3.11 : Durée moyenne de « pbloc » versus « pblocsy » chez les sujets LT

	Occurrence	Durée	Ecart-type
pbloc	489	0,48	0,33
pblocsy	184	1,09	0,59

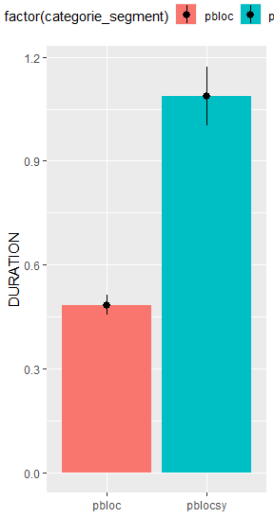

Figure 3.10 : Sujets LT : durée des pbloc (rouge) et des pblocsy (bleu)

Les pauses de blocages simples étaient de 0,48 secondes en moyenne et les pauses de blocages à fonction syntaxique étaient de 1,09 seconde en moyenne. Un test de Wilcoxon a été réalisé et confirme que les pauses de blocage à fonction syntaxiques sont plus longues que les pauses de blocage simples ($p < 2{,}2^{e\text{-}16}$).

Résultats individuels par locuteur

Une comparaison de la durée des pauses de blocage par rapport à celle des pauses de blocage à fonction syntaxique a aussi été réalisée individuellement pour les sujets LT. Le ratio durée pbloc/durée pblocsy semble varier de manière importante entre les sujets.

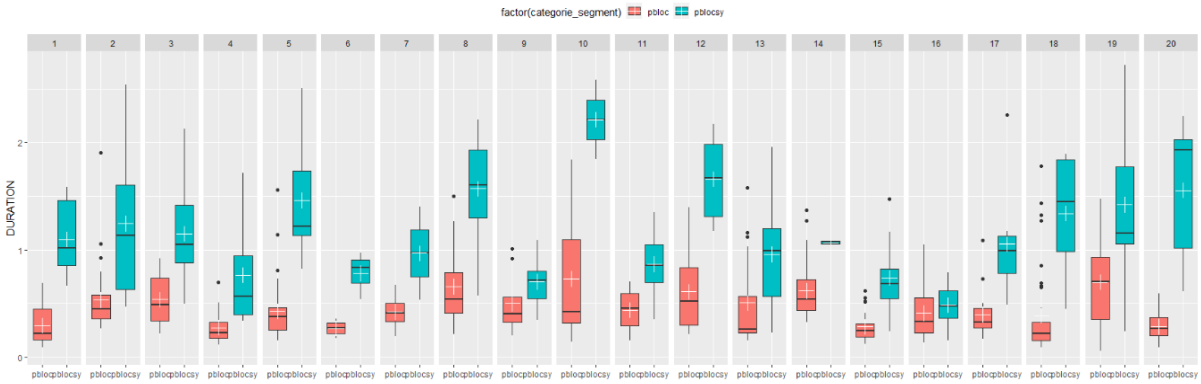

Figure 3.11 : durée des pbloc (rouge) et des pblocsy (bleu) par sujet.

Comparaison de la durée des pauses d'injection simples vs pauses d'injection à fonction syntaxique.

Nous avons comparé la durée des pauses d'injection simples (pinj) à celle des pauses d'injection à fonction syntaxique (pinjy).

Tableau 3.12 : Durée moyenne de « pinj » versus « pinjsy » chez les sujets LT

	Occurrence	Moyenne durée pause	Ecart-type
pinj	29	0,36	0,16
pinjsy	5	0,67	0,17

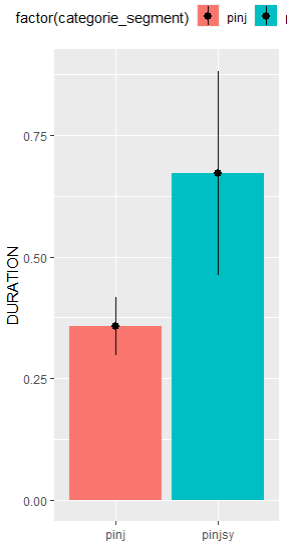

Figure 3.12 : Sujets LT : durée de pinj (rouge) et des pinjsy (bleu)

Les pauses d'injection simples étaient de 0,36 secondes en moyenne et les pauses d'injection à fonction syntaxique étaient de 0,67 secondes en moyenne. Un test de Wilcoxon a été réalisé et confirme que les pauses d'injection à fonction syntaxique sont plus longues que les pauses d'injection simples (p = 0,002329).

Il est important de noter que seules 5 pinjsy ont été retrouvées dans le discours de nos sujets utilisant la voix oro-œsophagienne.

Comparaison de la durée des pauses de blocage simple vs pauses de blocage à fonction d'hésitation.

Nous avons comparé la durée des pauses de blocage simple (pbloc) à celle des pauses de blocage à fonction d'hésitation (pbloches).

Tableau 3.13 : Durée moyenne de « pbloc » versus « pbloches » chez les sujets LT

	Occurrence	Moyenne durée pause	Ecart-type
pbloc	489	0,48	0,33
pbloches	41	1,01	0,63

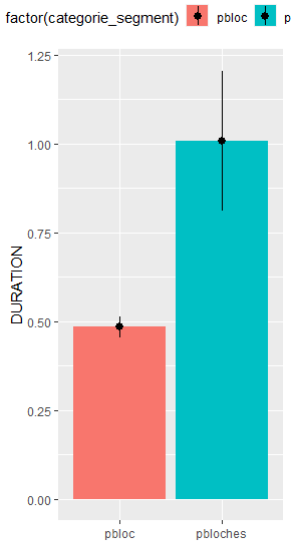

Figure 3.13 : Sujets LT : durée des pbloc (rouge) et des ploches (bleu)

Les pauses de blocage simples étaient de 0,48 secondes en moyenne et les pauses de blocage à fonction d'hésitation étaient de 1,01 seconde en moyenne. Un test de Wilcoxon a été réalisé et confirme que les pauses de blocage à fonction d'hésitation sont plus longues que les pauses de blocage simples ($p = 2{,}359^{e\text{-}09}$).

Comparaison de la durée des pauses de blocage à fonction syntaxique vs pauses de blocage à fonction d'hésitation.

Nous avons comparé la durée des pauses de blocage à fonction syntaxique (pblocsy) à celle des pauses de blocage à fonction d'hésitation (ploches).

Tableau 3.14 : Durée moyenne de « pblocsy » versus « ploches » chez les sujets LT

	Occurrence	Moyenne durée pause	Ecart-type
ploches	41	1,01	0,63
pblocsy	184	1,09	0,59

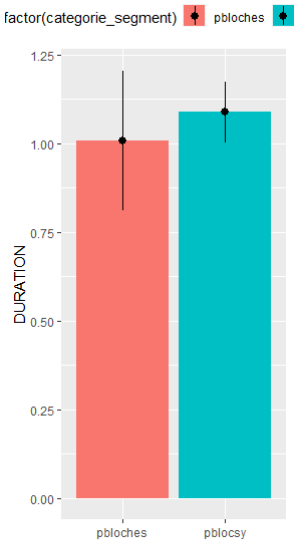

Figure 3.14 : Sujets LT : durée des ploches (rouge) et des pblocsy (bleu)

La durée des pauses de blocage syntaxiques est équivalente à celle des pauses de blocage d'hésitation. Le test de Wilcoxon réalisé ne retrouve pas de différence significative entre les deux types de pauses (p = 0,3748).

Comparaison de la durée des pauses d'hésitation remplies chez les sujets LT vs CTRL.
Nous avons étudié les pauses à fonction d'hésitation remplies en comparant leur durée chez les sujets CTRL (nommés sujets sains dans le graphique) et chez les sujets LT.

Tableau 3.15 : durée moyenne des phesr chez les deux groupes de sujets

	Occurrence	Moyenne durée pause	Ecart-type
LT	29	0,35	0,21
CTRL	161	0,52	0,31

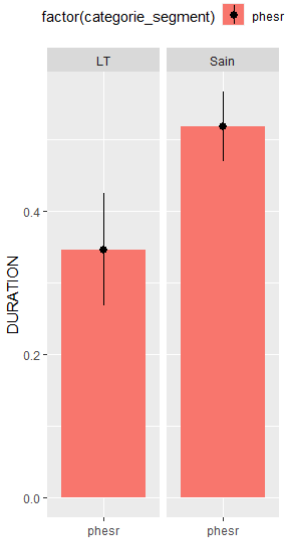

Figure 3.15 : durée moyenne des phesr chez les sujets LT (à gauche) et chez les sujets sains (à droite)

La durée des pauses d'hésitation remplies est plus importante chez les sujets sains (0,52 secondes) que chez les sujets LT (0,35). Un test de Wilcoxon a été effectué et retrouve une différence significative entre les deux groupes de sujets (p = 0.001146).

3.2 – Durée des syllabes

3.2.1 - Comparaison de l'allongement final aux frontières prosodiques

L'objectif était d'étudier la capacité des sujets à réaliser un allongement de la dernière syllabe, précédant une frontière prosodique. Pour cela nous avons comparé la durée des syllabes situées dans ces positions à celle des syllabes normales. Cette comparaison a été effectuée au sein de chaque groupe, puis entre nos deux populations.

Tableau 3.16 : durée des syllabes et syllabes finales chez sujets LT et CTRL.

	Quantité	Durée moyenne (en s)	Ecart-type
Syllabes LT	2096	0,22	0,11
Syllabes CTRL	3499	0,20	0,13
Finale LT	286	0,34	0,13
Finale CTRL	284	0,38	0,14

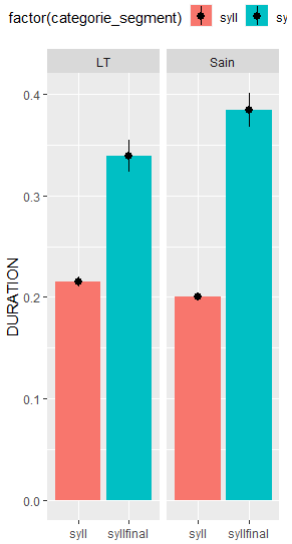

Figure 3.16 : durée moyenne des syllabes (rouge) et des syllabes finales (bleu) chez les sujets LT (à gauche) et chez les sujets sains (à droite).

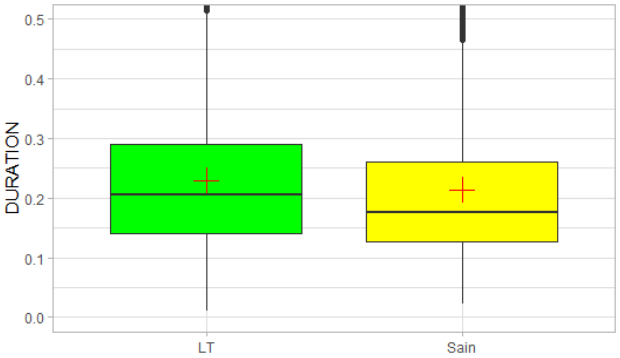

Figure 3.17 : durée moyenne des syllabes (toutes syllabes confondues) chez les sujets LT (vert) et chez les sujets sains (jaune).

Nous avons utilisé un modèle mixte, comprenant la catégorie du segment - ici syllabe et syllabe finale - et la durée de ce segment, selon le groupe. On trouve alors un triple effet :

- L'effet de catégorie du segment, qui est significatif (p<0,0001) : les syllabes finales sont plus longues que les autres syllabes du discours semi-spontané, que ce soit chez les sujets LT ou les sujets CTRL (nommés sujets sains dans les graphiques).
- L'effet de groupe est lui non significatif (p = 0,4009) : la durée des syllabes (toutes syllabes confondues) est donc équivalente chez les sujets LT et sains (voir figure 3.17).
- L'effet de l'interaction catégorie segment / groupe, qui est significatif (p<0,0001) : l'allongement final est plus important chez les sujets sains que chez les sujets LT.

3.3. Débit de parole

Nous avons comparé le débit de parole des sujets sains et des sujets opérés d'une laryngectomie totale. Pour rappel le débit de parole est le nombre de syllabes/seconde en incluant dans le calcul les pauses du discours.

Tableau 3.16 : Débit de parole moyen chez les sujets LT et CTRL.

	Nombre de sujets	Débit de parole moyen	Ecart-type
LT	20	2,12	0,69
CTRL	21	3,20	0,57

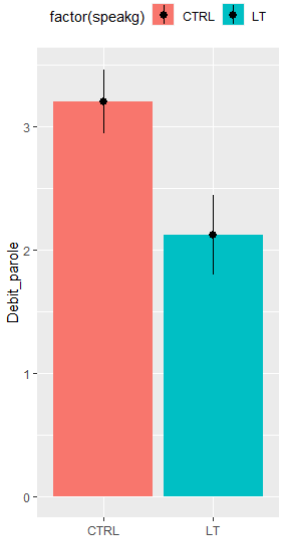

Figure 3.17 : débit de parole chez les sujets CTRL (rouge) et chez les sujets LT (bleu)

Le débit de parole moyen des sujets opérés d'une laryngectomie totale est de 2,12 syllabes/seconde. Le débit de parole moyen des sujets sains est de 3,2 syllabes/secondes. Le débit de parole des sujets laryngectomisés est plus lent que celui des sujets sains.

Un test de Wilcoxon a été réalisé et retrouve une différence significative entre les débits de parole moyens des deux groupes de sujets ($p= 3,595^{e-06}$)

3.4. Vitesse articulatoire

Nous avons comparé la vitesse articulatoire des sujets opérés d'une laryngectomie totale avec celle des sujets sains. Pour rappel, la vitesse articulatoire est le nombre de syllabes/seconde, en dehors des temps de pause du discours.

Tableau 3.17 : Vitesse articulatoire moyenne chez les sujets LT et CTRL

	Nombre de sujets	Vitesse articulatoire moyenne	Ecart-type
LT	20	4,43	0,77
CTRL	21	4,61	0,77

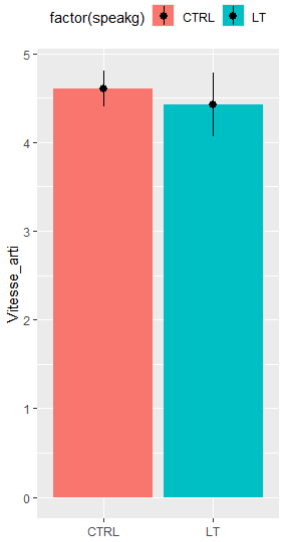

Figure 3.18 : Vitesse articulatoire chez les sujets CTRL (rouge) et chez les sujets LT (bleu).

La vitesse articulatoire moyenne des sujets laryngectomisés est de 4,43 syllabes/seconde. La vitesse articulatoire moyenne des sujets sains est de 4,61 syllabes/seconde. La vitesse articulatoire des deux groupes de sujets est équivalente.

Un test de Wilcoxon a été réalisé et ne retrouve pas de différence significative entre nos deux groupes de sujets (p= 0,2562).

4. Discussion

Notre travail avait pour but d'étudier le rythme dans la parole oro-œsophagienne, et plus particulièrement la distribution des pauses au sein de celle-ci par rapport à celle de sujets sains. En effet, les pauses représentent une des composantes du rythme, nécessaires à la réalisation de diverses fonctions essentielles à la parole. Toutefois, il s'agit d'un équilibre précis entre parole et silence, qui semble être bousculé chez les patients laryngectomisés du fait du besoin physiologique de blocage que leur demande l'utilisation de la voix oro-œsophagienne. Pour cette recherche expérimentale, nous avions alors émis deux hypothèses en nous appuyant sur les résultats de la littérature (Bellandese et al., 2001; Galant, 2019; Moerman et al., 2015; Most et al., 2000; Motta et al., 2001; van Rossum, 2005; van Rossum et al., 2008) : la première étant que la parole oro-œsophagienne nécessiterait et induirait donc plus de pauses que la parole saine ; la seconde que ces pauses se distingueraient selon leur fonction dans le discours par une différence de durée.

L'objectif principal de ce mémoire consistait donc en une description phonétique de l'organisation temporelle de la parole oro-œsophagienne, qui pourrait à long terme permettre une adaptation du bilan orthophonique des patients ayant bénéficié d'une laryngectomie totale. L'analyse précise et objective du rythme de la parole servirait alors à développer des protocoles de rééducation développant l'axe rythmique de la parole dans la prise en charge des sujets oro-œsophagiens.

Tous types confondus, nos résultats nous permettent d'observer un nombre de pauses plus important chez les sujets laryngectomisés que chez les sujets sains. En effet, nous obtenons des résultats significatifs ($p = 0,0007261$) quant à la différence de quantité de pauses entre nos deux groupes de sujets. Nos résultats rejoignent alors ceux de diverses études sur la parole oro-œsophagienne, indiquant que la quantité de pauses chez le sujet oro-œsophagien est plus importante que chez le sujet sain (Bellandese et al., 2001 ; Moerman et al., 2015 ; Most et al., 2000 ; Motta et al., 2001). Dans son travail de thèse, Camille Galant a également démontré que les sujets oro-œsophagiens produisent plus de pauses que les sujets sains dans une tâche de lecture (Galant, 2019). De multiples raisons provoquent chez le sujet utilisant cette voix de substitution l'ajout de nombreuses pauses et donc une altération importante du rythme.

Premièrement, à la suite d'une laryngectomie totale, la phonation laryngée étant remplacée par l'érygmophonie, le réservoir œsophagien devient la source d'air phonatoire. Or sa capacité est plus faible que le réservoir pulmonaire des sujets sains, ce qui force les personnes laryngectomisées à produire davantage de pauses de blocage afin de le remplir (Most et al., 2000; Motta et al., 2001; Ng, 2011). Cet ajout de pauses de blocage est aussi, de fait, nécessaire à la sonorisation. L'hyperarticulation évoquée dans notre état des lieux, mise

en évidence par Motta, provoque également la réalisation de pauses supplémentaires. Même si le but premier est d'améliorer l'intelligibilité, cela perturbe le rythme (Motta et al., 2001). De plus, plusieurs auteurs ont corrélé une diminution de l'intelligibilité à la présence d'un nombre plus important de pauses chez le sujet oro-œsophagien (Bellandese et al., 2001; Moerman et al., 2015; Motta et al., 2001). De fait, si l'on tend à penser que la plainte principale des patients utilisant une phonation oro-œsophagienne réside dans l'altération de leur voix, les nombreuses pauses adjointes à leur parole et leur impact sur le rythme et l'intelligibilité semblent venir contredire cette idée. Si l'on associe cela aux études sur le VHI ayant montré une certaine indulgence des patients alaryngés envers leur nouvelle voix, nous pouvons émettre l'hypothèse que la gêne des patients concerne finalement plus la parole que la voix. Cela serait toutefois à confirmer au cours d'une autre étude incluant une analyse des auto-évaluations des patients.

Notre hypothèse concernant le nombre plus important de pauses chez les sujets oro-œsophagiens par rapport aux sujets sains est donc validée, à la fois par notre analyse des données de la littérature et notre expérience, réalisée en discours semi-spontané.

Cette hypothèse est d'ailleurs également émise et confirmée dans le cadre d'autres pathologies – bégaiement, dyslexie, maladie de Parkinson, d'Alzheimer, etc. –, elles aussi génératrices de pauses venant modifier la structure rythmique du discours que l'on retrouve dans la parole saine (Barkat-Defradas et al., 2009; Smith & Caplan, 2018; Teixeira et al., 2012; Lalain, 2012 ; Yunusova et al., 2016).

Nous avons par la suite voulu étudier les pauses de manière plus précise. Les différentes fonctions qu'elles peuvent revêtir dans les différents types de parole, leur quantité et leur durée. Notre hypothèse étant pour rappel que ces fonctions se distingueraient en termes de durée dans le discours.

Le caractère plus spontané que présente une tâche de description d'image, comparé à la lecture, permet une plus grande liberté dans la façon de s'exprimer. Toutefois, cela induit aussi un besoin plus fort de planification, à la fois cognitive et physiologique. En lecture, le sujet n'a pas à réfléchir au contenu de son discours et les pauses sont presque toutes programmées par la ponctuation. De fait, nous comptabilisons en tâche de description d'image 17 types de pauses différents, soit plus qu'en lecture (seulement 10). Huit sont similaires à ceux trouvés dans cette tâche : des pauses à visée syntaxique, seules pour les sains, avec blocage chez les patients laryngectomisés ; des pauses de blocage pour les sujets alaryngés, avec ou sans fonction syntaxique associée ; des pauses de respiration ou encore des pauses de focalisation et d'autocorrection. Nous excluons la pause de décodage, qui n'avait pas lieu d'être dans une tâche de description d'image, et la majorité des types de pauses que nous avons ajoutés se révèlent être des co-occurrences de fonctions pour une

même pause (pblocsy, prespihes, prespisy, etc…). On observe également l'apparition de pauses à fonction d'hésitation.

Cette différence peut être expliquée par le type de discours utilisé, qui n'implique que très peu de contraintes au locuteur en comparaison à la lecture. Nous l'avons vu, en parole normale, notre temps de parole serait à 40 voire 50% composé de pauses durant un description d'image (Goldman-Eisler, 1968 in Hirsch et al., 2018). Celles-ci sont nécessaires à la fois pour la planification des énoncés par les locuteurs, et pour le traitement par les auditeurs (Campione & Véronis, 2005) ; et d'après Lalain et al., elles apporteraient autant au message que la parole elle-même (Lalain et al., 2016).

De plus, cette tâche descriptive demande plus de charge cognitive, et différents types de pauses apparaissent. Les pauses isolées sont prédominantes dans les situations préparées (journaux parlés, homélies, discours politiques, etc…) et plus rares dans les situations non préparées (conversation informelle, narration conversationnelle, etc…) (Grosman et al., 2018). Aussi, une fonction est difficilement attribuable de manière certaine à une pause précise (Goldman et al., 2010; Grosjean & Deschamps, 1975; Lalain et al., 2016). Il n'est donc pas étonnant de retrouver dans notre corpus des chevauchements de fonctions au sein d'une même pause en plus grand nombre, à la fois chez nos sujets sains et laryngectomisés.

Par la suite, Camille Galant a fait le choix de n'étudier que les pauses de blocage et les pauses de blocage à fonction syntaxique, car elles constituaient les pauses majoritaires de son corpus, paramètres influencés par la tâche de lecture, et plus particulièrement par la présence de ponctuation. Pour obtenir un traitement plus précis de nos données nous n'avons pas pu faire ce choix, le discours semi-spontané induisant plus de types de pauses.

Nous avons alors dans un premier temps procédé à ces comparaisons dans le cadre de la tâche de description d'image. Toutes pauses à fonction syntaxique confondues, nous obtenons une quantité de pauses à fonction syntaxique équivalente chez les sujets laryngectomisés et chez les sujets sains (avec un test de Wilcoxon montrant une différence non-significative : $p = 0.454$).

Ces résultats viennent appuyer ce qui avait été démontré en lecture : si les parleurs utilisant une voix oro-œsophagienne procèdent à beaucoup plus de pauses physiologiques que les sujets sains, le nombre de pauses syntaxiques est quant à lui presque identique dans les deux populations (Galant, 2019). La seule différence est que celles réalisées par les sujets utilisant la voix oro-œsophagienne sont accompagnées d'un blocage physiologique.

Cette similitude peut être attribuée, en lecture, à un respect par les sujets utilisateurs de voix oro-œsophagienne des frontières syntaxiques imposées par le texte (Galant, 2019). Or, la description est certes en partie dirigée par l'image choisie, qui donnent des éléments

clés à introduire dans leur discours, mais ils doivent en construire l'expression : du choix sémantique à la tournure syntaxique, tout en gérant leurs besoins physiologiques. Nous pouvons alors supposer que les sujets alaryngés tentent, malgré l'adjonction de nombreuses pauses physiologiques, de garder une structure syntaxique cohérente à celle retrouvée dans la parole saine.

Nous avons également comparé chez le sujet laryngectomisé la quantité de pauses de blocage simple par rapport à celle de pauses de blocage à fonction syntaxique. Chez nos sujets LT, les pauses physiologiques sont les pauses de blocage. Or, dans un souci d'efficacité, ils profitent des pauses syntaxiques pour bloquer de l'air. Il a donc été compliqué de dissocier syntaxique et physiologique chez les sujets LT. Toutefois, nous obtenons une différence significative ($p = 0,0004786$), avec une production plus importante de pauses de blocage simples (24,45 en moyenne) que de blocages à fonction syntaxique (9,65 en moyenne). Camille Galant avait elle aussi trouvé une différence significative entre ces deux types de pauses dans son analyse en lecture (Galant, 2019). Cela confirme le fait que, même si le sujet LT essaie d'optimiser les pauses à fonction syntaxique en y accolant un blocage, cela ne suffit pas à compenser le besoin physiologique important qu'induit la voix oro-œsophagienne. Il produit donc également de nombreux blocages au sein de la phrase, que l'on peut attribuer au besoin de remplissage du réservoir œsophagien dans l'objectif de sonoriser. La pause de blocage simple est d'ailleurs le type de pauses le plus représenté dans notre échantillon de parole semi-spontanée.

La lecture laisse moins de place à la planification que des tâches de description d'image ou de discussion spontanée. Notre corpus comprenait 291 pauses à fonction d'hésitation, tandis qu'aucune n'est retrouvée dans le corpus de lecture de Camille Galant, que ce soit chez les sujets sains ou laryngectomisés. Cela concorde avec les différentes études menées en parole normale, qui ne retrouvent des pauses d'hésitation que dans le discours spontané ou semi-spontané (Lalain et al., 2016). De la même manière, Grosman et al. et Candea retrouvent en discours non préparé plus de pauses non-structurantes que dans la lecture (pause d'hésitation – vide ou pleine – d'autocorrection, etc…) (Grosman et al., 2018 ; Candea, 2000). En lecture, nous procédons en effet à des pauses de décodage du texte plutôt qu'à des pauses d'hésitation.

Nous avons alors voulu comparer la quantité de ce type de pauses chez le sujet LT et le sujet CTRL, et toutes pauses à fonction d'hésitation confondues (phes, phesr, pbloches, prespihes, pinjhes), les sujets sains produisent significativement plus de pauses d'hésitation que les sujets LT ($p = 0,0002765$). Ils procèdent en effet à plus du double de pauses d'hésitation (11,57) par rapport aux sujets laryngectomisés (4,70). Nous ne nous attendions pas à une si grande différence étant donné que la description d'images demande autant, voire plus de planification aux sujets laryngectomisés du fait de la complexité de réalisation de la parole

oro-œsophagienne. Cette différence pourrait être due au fait que les nombreuses pauses de blocage simples dans la parole oro-œsophagienne rendent difficile l'interprétation de la présence ou non d'une hésitation. Sachant que les sujets laryngectomisés adaptent au maximum leur parole et que la priorité est donnée à la sonorisation, le sujet LT peut profiter de ses temps de pause de blocage pour planifier à la fois sa sonorisation, son articulation et son discours.

Nous avons ensuite étudié la quantité et la durée des pauses d'hésitation remplies (production d'un « euh », « hum », allongement important d'une syllabe, etc…) réalisées par nos deux populations. Les sujets sains produisent à nouveau beaucoup plus de pauses remplies (7,66 pauses en moyenne) que les sujets laryngectomisés (1,45 pause en moyenne) (p= 2,732 e-0,5). De la même manière, celles-ci sont significativement plus longues chez les sujets sains (p= 0,001146). Ces résultats sont cohérents avec la difficulté qu'ont les sujets oro-œsophagiens à produire une sonorisation. Il est donc peu pertinent et efficace pour eux de voiser un temps qui ne transmet pas de message verbal. Il est toutefois important de noter, de manière qualitative, que certains sujets LT produisent des pauses d'hésitation remplies. Celles-ci sont certes plus courtes et moins nombreuses que celle produites par les sujets contrôles, mais cela reste une prouesse au vu de la complexité que représente la réalisation de la parole oro-œsophagienne. On observe d'ailleurs, qualitativement, que les sujets produisant des pauses d'hésitation remplies présentent généralement un discours assez fluide et naturel. Le fait de se rapprocher du mode de fonctionnement des sujets sains pourrait donc induire une meilleure perception et montrer une meilleure maîtrise de la voix oro-œsophagienne. Les pauses remplies sont les diffluences les plus courantes en français (Grosjean & Deschamps, 1975). Elles seraient très importantes dans le discours, permettant de ne pas perdre le tour de parole le temps de planifier notre message (Laroche-Bouvy, 1984).

A l'inverse, nos résultats concernant la comparaison de quantité de pauses d'auto-correction démontrent que leur quantité moyenne est équivalente dans nos deux groupes de sujets (p= 0,175). La pause d'auto-correction a pour rôle dans le discours de pallier un mauvais départ ou un accroc (Nespoulous, 2014 in Dodane & Hirsch, 2018). On observe le plus souvent un changement de mots ou une rectification de la prononciation. Ces résultats peuvent être le reflet de la charge cognitive accompagnant la tâche de discours semi-spontané, que le sujet ait une pathologie ou non.

L'étude des pauses en parole normale a mis en évidence des variations de leur durée : courtes (<200ms), moyennes (200-1000ms) ou longues (>1000ms) (Campione & Véronis, 2002). Plusieurs auteurs les ont analysées et ont trouvé différents facteurs expliquant ces variations, dont principalement la fonction de la pause et le type de parole. Toutefois, aucun consensus clair n'a pu être trouvé.

Nos résultats concordent avec notre hypothèse selon laquelle les pauses auraient une durée différente selon leur fonction dans la phrase, et sont en adéquation avec les diverses études de la parole normale présentes dans la littérature.

En effet, on retrouve dans notre analyse une différence significative de durée entre la pause de blocage à fonction syntaxique et la pause de blocage « simple » ($p < 2,2e-16$), la première étant deux fois plus longue (1 seconde) que la seconde (0,5 secondes). Cela montre que les sujets laryngectomisés respectent la structure syntaxique du discours en s'adaptant aux frontières prosodiques. Nous pouvons supposer que cette stratégie leur permettrait de dissimuler les pauses de blocage simples : réduire leur durée au maximum aurait pour objectif de perturber le moins possible la structure du discours, et donner une durée plus longue aux pauses syntaxiques permettrait de mettre en exergue la structure syntaxique de leur phrase, capitale pour la bonne compréhension du message par l'interlocuteur. Cette différence de durée étaient déjà présente en lecture, dans la thèse de Camille Galant. Les fonctions corrélées à une durée plus importante sont majoritairement la fonction syntaxique et la fonction physiologique de respiration (Goldman et al., 2010 ; Lalain et al., 2012). Dans une étude récente, Grosman et al. décrivent d'ailleurs la localisation syntaxique comme le paramètre influençant le plus la durée des pauses (Grosman et al., 2018).

De plus, la cooccurrence des fonctions sur une même pause semble engendrer une durée supérieure à celle d'une pause « simple ». On trouve des résultats significatifs chez les sujets LT : une durée plus importante est mise en avant pour les pblocsy par rapport aux pbloc ($p < 2,2e-16$), pour les pinjsy par rapport aux pinj ($P=0,002329$) et pour les pbloches par rapport aux pbloc ($p = 2,359e-09$). A l'inverse, on ne constate pas de différence de durée lorsque l'on compare des pauses simples entre elles ou des pauses hybrides entre elles, ce qui nous conforte dans cette idée. En effet, aucune différence significative n'est retrouvée entre pause d'injection simple et pause de blocage simple ($p = 0,141$), entre pause de blocage syntaxique et pause d'injection syntaxique ($p = 0,09098$) et entre pause de blocage syntaxique et pause de blocage à fonction d'hésitation ($p = 0,3748$). Toutefois, les résultats en rapport avec les pauses d'injection sont à prendre avec précaution, car on ne les retrouve qu'en faible nombre au sein de notre corpus (29 pauses d'injection simples pour 5 pauses d'injection syntaxiques), et seulement 4 de nos 20 sujets laryngectomisés en produisent. Cela rejoint néanmoins les observations faites en parole normale : une étude avait prouvé que les pauses syntaxiques respiratoires sont significativement plus longues que les pauses syntaxiques simples dans toutes les tâches étudiées (lecture, description, discussion) (Lalain et al., 2016). Hirsch remarque et précise également que sans événements physiologiques, les pauses syntaxiques simples sont les plus courtes de leur corpus (Hirsch et al., 2018).

Nous avons ensuite voulu approfondir notre analyse rythmique via une étude de l'allongement final prosodique, ainsi que la comparaison du débit et de la vitesse articulatoire des deux groupes de sujets.

Concernant cet allongement final, nos résultats rejoignent ceux de Camille Galant. En effet, que ce soit chez les sujets sains ou les sujets laryngectomisés, les syllabes finales précédant une pause syntaxique sont significativement plus longues que le reste des syllabes du discours semi-spontané ($p<0,0001$). Dans une situation de description d'image, les sujets LT produisent donc également cet allongement de la dernière syllabe. Celui-ci permet globalement de mettre en évidence la structure syntaxique de leurs phrases, et est caractéristique du rythme de la parole en français. Or, dans ses études de 2005 et 2008, Van Rossum statue que cet allongement prosodique est remplacé par une pause chez le sujet LT. Toutefois, celles-ci n'analysaient que la langue néerlandaise (van Rossum, 2005, 2008). La présence d'un allongement de la dernière syllabe dans la langue française n'est plus à prouver en parole normale. Nos sujets ayant bénéficié d'une laryngectomie totale font preuve depuis leur opération d'une grande capacité d'adaptation, et essaient de retranscrire ces habitudes phonétiques et rythmiques dans leur nouvelle parole. La présence d'un allongement des syllabes finales en lecture (Galant, 2019), et en parole semi-spontanée, dans notre étude, montre à quel point certaines structures rythmiques sont ancrées dans notre parole, si bien qu'elles en deviennent capitales, même chez des voix très déstructurées.

Nous obtenons également des résultats cohérents en ce qui concerne la longueur de ces syllabes finales, étant données les capacités phonatoires de nos sujets laryngectomisés, comparé à des sujets sains. Comme nous l'attendions, ces derniers produisent des syllabes finales significativement plus longues que les sujets LT ($p<0,0001$). Toutefois, lorsque l'on étudie la durée des toutes les syllabes confondues, nous ne retrouvons pas de différence significative entre sujets CTRL et sujets LT ($p = 0,4009$). Les sujets utilisant une voix oro-œsophagienne produisent des syllabes (hors syllabes finales) légèrement plus longues que les sujets contrôles, ce qui compense la différence de durée entre leurs syllabes finales. Cette équivalence de durée des syllabes entre LT et sains permet d'expliquer les résultats retrouvés en termes de vitesse articulatoire présentés ci-après.

Notre étude portant sur l'organisation rythmique de la parole oro-œsophagienne, nous avons voulu comparer le débit de parole chez les sujets contrôles et chez les sujets laryngectomisés. Pour rappel le débit de parole correspond au nombre de syllabes produites par seconde, en prenant en compte les temps de pause du discours. Nous obtenons des résultats significatifs : un débit de parole moyen plus lent est observé chez les sujets LT comparé aux sujets sains. Les sujets opérés d'une laryngectomie totale produisent 2,12 syllabes/seconde, contre 3,2 pour les sujets sains ($p= 3,595e-06$). Nos résultats rejoignent ceux de Camille Galant en lecture. Les complications inhérentes à la production de la parole oro-

œsophagienne, accompagnée de pauses plus nombreuses et plus longues, provoquent de manière logique un ralentissement du débit chez les sujets LT.

Afin d'évaluer plus précisément le timing de la parole oro-œsophagienne et l'impact des pauses sur celle-ci, nous avons alors calculé la vitesse articulatoire. Cette dernière représente le nombre de syllabes émises par secondes, cette fois sans prendre en compte les temps de pauses. De nombreux auteurs décrivent que l'ablation de l'os hyoïde lors de la laryngectomie totale provoque une réduction de la force et de la mobilité linguale, impactant donc la précision et la durée de l'articulation (Bellandese et al., 2001 ; Gandour & Weinberg, 1980 ; Ng, 2019). Connaissant donc les difficultés de production articulatoire de nos patients laryngectomisés, nous nous attendions comme dans l'étude de Camille Galant à une valeur moins élevée chez les sujets LT. Or, nous obtenons des résultats équivalents entre nos deux groupes de sujets, avec une vitesse articulatoire de 4,43 syllabes/seconde chez les LT et 4,61 syllabes/seconde chez les CTRL. D'après le test de Wilcoxon effectué, cette différence n'est pas significative (p=0,2562).

Nous pouvons alors émettre plusieurs hypothèses. La première concerne le type de discours étudié : nos sujets laryngectomisés parlent peut-être plus rapidement en semi-spontané qu'en lecture. Une comparaison avec un discours spontané serait donc intéressante pour infirmer ou confirmer nos résultats. Une comparaison lecture/semi-spontané chez nos sujets LT, ainsi qu'une analyse des valeurs du cepstral peak proeminence (CPP, qui est une mesure objective et robuste du timbre de la voix, reflétant la quantité d'harmoniques sans avoir besoin nécessairement d'un signal périodique) et des mesures de l'intelligibilité seraient aussi à envisager. Cela nous permettrait en effet de vérifier si les sujets laryngectomisés ne parleraient pas plus vite en semi-spontané, au détriment de la qualité de leur voisement et de leur intelligibilité.

Une autre hypothèse serait que les sujets sains parleraient plus lentement en discours semi-spontané, du fait de la charge cognitive plus importante que la situation de description d'image implique comparée à la lecture. Pour nos sujets laryngectomisés, nous supposons que les capacités d'adaptation qu'ils ont dû mettre en place au quotidien leur permettraient d'amoindrir l'impact de la charge cognitive de cette tâche sur leur parole, mais aussi que les nombreuses pauses physiologiques de blocages adjointes à leur parole pourraient également servir à la planification cognitive, et cette dernière ne ralentirait donc pas l'enchaînement articulatoire.

Enfin, ces résultats concordent avec la comparaison des durées de syllabes effectuée précédemment : les sujets alaryngés et les sujets sains présentent des durées de syllabes presque identiques (220 ms en moyenne pour les sujets LT et 200 ms pour les sujets sains).

5. Limites et perspectives

La recherche sur la parole oro-œsophagienne a été peu prolifique jusqu'alors. Ce domaine commence à se développer face à la nécessité de répondre toujours plus efficacement et écologiquement aux besoins de nos patients et, dans un souci de prise en charge globale, différents aspects touchant cette pathologie mériteraient d'être plus profondément explorés. Ce travail de mémoire, qui faisait suite à la thèse de Camille Galant, avait pour objectif d'explorer l'aspect rythmique de cette parole en condition semi-spontanée. Nos résultats sont encourageants, mais ils comportent un certain nombre de limites qu'il serait judicieux d'exposer afin de pouvoir y remédier par la suite, et ouvrent d'autres perspectives d'étude de la parole oro-œsophagienne.

Le faible nombre de sujets constitue une des limites principales de notre recherche. Les résultats seraient d'autant plus fiables et représentatifs s'ils provenaient d'un plus grand nombre de sujets sains et de sujets laryngectomisés. Il serait intéressant d'envisager ce protocole de recherche avec un échantillon de population plus conséquent.

La seconde limite que nous pouvons évoquer est l'absence de comparaison de nos résultats avec les résultats d'une évaluation perceptive de nos sujets et/ou les résultats d'une auto-évaluation. La subjectivité de ce type d'évaluations s'accompagne de nombreuses limites, mais cela nous aurait permis de faire des corrélations entre ce que nous avons observé et ce qui en transparaît dans l'expression et la réception du message produit par nos sujets. L'impact de ce trouble sur le ressenti des sujets eux-mêmes aurait également pu être observé grâce à celles-ci.

De plus, la comparaison de nos résultats avec ceux obtenus en lecture par Camille Galant dans sa thèse n'a pu être que qualitative. En effet, nous n'avons pas pu procéder à une analyse statistique comparant les deux types de discours.

Notre corpus constitue aussi une limite, dans le sens où nous n'avons étudié qu'un seul type de discours : le semi-spontané. Le support de la description d'image induit un certain cadre et des contraintes thématiques et discursives. La personne n'est pas totalement libre dans son discours, et nous estimons que ces contraintes soulageraient plus ou moins la planification du discours. Dans une étude ultérieure, notre expérimentation sera réalisée avec cette fois-ci des échantillons de parole spontanée. Puisque la parole des sujets LT n'aura pas le soutien du support imagé (description), il est alors envisageable que la distribution des pauses et de leur durée soit remaniée. Nous nous questionnons sur le fait que ce type de discours nécessiterait peut-être plus de pauses à fonction cognitive de type hésitation, planification, auto-correction, et donc que les sujets LT auraient ici la nécessité d'optimiser leurs blocages selon des contraintes de planification du discours.

Afin de valider nos hypothèses au sujet des résultats obtenus pour la vitesse articulatoire, il serait également intéressant de procéder à une comparaison de cette dernière chez des sujets sains et des sujets laryngectomisés, pris de manière individuelle, en lecture par rapport à celle obtenue en semi-spontané.

Enfin, le discours semi-spontané est un biais important dans notre étude, car les échantillons de parole ne sont pas égaux en durée. Un temps de parole plus long implique plus d'occasions de pause dans le discours.

Outre les évaluations subjectives et l'étude du rythme en discours spontané, il paraît intéressant d'étudier aussi un autre aspect majeur de la transmission d'un message dans une situation de communication : la prosodie. En effet, tout comme le rythme, les contours intonatifs pourraient être fortement impactés par l'ajout de nombreuses pauses, nécessaires aux sujets laryngectomisés, et par leur durée. Peu d'études de la prosodie oro-œsophagienne existent (van Rossum, 2005; van Rossum et al., 2002, 2008 ; Cuenca & Barrio, 2010) et aucune n'a étudié la prosodie du français.

6. Conclusion

Nous avons donc, dans ce travail, analysé la parole oro-œsophagienne. Notre but était de mettre en lumière les difficultés rythmiques engendrées par la perte de l'organe phonatoire, venant s'ajouter aux différentes altérations acoustiques. Nous l'avions vu, la littérature à ce sujet ne faisait état que d'une quantité plus importante de pauses dans le discours des patients laryngectomisés, sans description plus poussée de ces dernières. Dans sa récente thèse, Camille Galant a proposé d'approfondir l'étude de cette parole sans larynx, en procédant à l'analyse des différentes pauses retrouvées dans le cadre d'une tâche de lecture. En sont ressorties différentes observations : les patients laryngectomisés produisent effectivement plus de pauses que les sujets contrôles, mais tenteraient de les placer de façon à respecter la structure syntaxique du discours. Elle témoigne également d'un allongement des syllabes finales situées aux frontières syntaxiques, contrairement aux autres études réalisées à ce sujet.

Notre travail de mémoire constitue donc la suite de cette expérimentation, s'intéressant à l'organisation rythmique et à la distribution des pauses dans la parole oro-œsophagienne, toujours comparées à celles des sujets sains, mais cette fois lors d'une tâche de description d'image. Nous avions deux hypothèses, qui semblent être confirmées par les différents résultats qu'ont amenés nos expérimentations. Dans un premier temps, et en cohérence avec les recherches précédentes, on observe bien une quantité de pauses plus importante chez nos sujets alaryngés que chez les sains. Cela est expliqué par le défi physiologique que représente l'utilisation de cette voix.

Toutefois, nous observons que, malgré les difficultés de production induites par la disparition de l'organe phonatoire, les patients opérés d'une laryngectomie totale tentent bel et bien de se rapprocher le plus possible de leur schéma de parole antérieur. Ils s'efforcent de respecter la structure syntaxique du discours, et essaient d'optimiser leurs pauses de différentes manières : raccourcir la durée des pauses physiologiques intra-phrases ou encore utiliser la cooccurrence de fonctions au sein d'une même pause afin d'en réduire le nombre. On note d'ailleurs, comme Camille Galant l'avait montré en lecture, un nombre équivalent de pauses à visée syntaxique chez nos deux groupes de sujets.

Cette tâche a révélé un plus grand nombre de types de pauses qu'en lecture, ainsi qu'une utilisation plus fréquente de pauses dites « hybrides ». Celles-ci se distinguent d'ailleurs en durée des pauses à fonction unique, ce qui vient confirmer notre seconde hypothèse : selon leur fonction dans le discours les pauses ont une durée différente.

Tout comme ils essaient de respecter la structure syntaxique de la parole, ils tentent également d'en reproduire la prosodie. Nous trouvons en effet, comme Camille Galant, un allongement prosodique des syllabes se trouvant aux frontières syntaxiques du discours, certes moins long que celui des sujets sains, mais tout de même présent et repérable. De plus, on observe également, chez les patients maîtrisant le mieux leur nouvelle voix, des pauses d'hésitation remplies.

Enfin, nous retrouvons, à l'instar de la tâche de lecture, un débit de parole plus lent chez les sujets alaryngés que chez les sujets contrôles, ce dernier étant impacté par l'adjonction de ces nombreuses pauses physiologiques. Mais, résultat étonnant, la vitesse articulatoire ne montre, elle, pas de différence significative.

Nos résultats, et les différentes analyses que nous avons pu en faire, montrent l'importance des pauses au sein de la parole, ainsi que l'impact qu'une désorganisation rythmique peut avoir sur son élaboration et sa production. Toutefois, nous avons pu déceler différentes limites au sein de notre expérimentation, impliquant un besoin de solidification et d'étayage des différents éléments mis en lumière. Une analyse similaire à la nôtre est en cours, étudiant cette fois-ci le discours spontané, afin de compléter l'analyse de tous les styles de parole et de permettre le croisement des données et résultats recueillis. Obtenir une description phonétique et rythmique de la parole oro-œsophagienne, objective et consensuelle, pourrait donner à la prise en charge orthophonique un nouvel angle, à la fois lors de l'évaluation mais aussi de la rééducation de ces patients.

Bibliographie

Allali, A., Huche, F. L., & Hilgers, F. (2008). *La voix sans larynx : Manuel d'apprentissage des voix oro- et trachéo-oesophagiennes à l'usage des laryngectomisés porteurs et non porteurs d'implant phonatoire et de leurs rééducateurs*. De Boeck Supérieur. https://books.google.fr/books?id=CYVLswEACAAJ

Allegra, E., La Mantia, I., Bianco, M. R., Drago, G. D., Le Fosse, M. C., Azzolina, A., Grillo, C., & Saita, V. (2019). Verbal performance of total laryngectomized patients rehabilitated with esophageal speech and tracheoesophageal speech : Impacts on patient quality of life. *Psychology Research and Behavior Management*, *12*, 675-681. https://doi.org/10.2147/PRBM.S212793

Antin, F. (2017). *Réhabilitation après laryngectomie totale : État des lieux et établissement d'un support d'éducation thérapeutique* (p. 81) [Thèse d'exercice : Médecine]. Angers.

Barkat-Defradas, M., Frederique Gayraud, Claire Cadilhac, & Hyeran Lee. (2009, décembre 4). *Pauses et hésitations dans le discours de patients Alzheimer et chez la personne âgée saine*. 3° Journées de Phonétique Clinique. https://halshs.archives-ouvertes.fr/halshs-00801855

Bellandese, M. H., Lerman, J. W., & Gilbert, H. R. (2001). An acoustic analysis of excellent female esophageal, tracheoesophageal, and laryngeal speakers. *Journal of Speech, Language, and Hearing Research: JSLHR*, *44*(6), 1315-1320. https://doi.org/10.1044/1092-4388(2001/102)

Bellandese, Mary H. (2009). Fundamental frequency and gender identification in standard esophageal and tracheoesophageal speakers. *Journal of Communication Disorders*, *42*(2), 89-99. https://doi.org/10.1016/j.jcomdis.2008.08.005

Boersma, P. (2001). Praat, a system for doing phonetics by computer. *Glot. Int.*, *5*(9/10), 341-345.

Borrie, S. A., Lansford, K. L., & Barrett, T. S. (2018). Understanding dysrhythmic speech : When rhythm does not matter and learning does not happen. *The Journal of the Acoustical Society of America*, *143*(5), EL379-EL385. https://doi.org/10.1121/1.5037620

Bretagne, E. (2008). Être laryngectomisé : Réhabilitation et réadaptation. *Psycho-Oncologie*, *2*(2), 83-85. https://doi.org/10.1007/s11839-008-0075-z

Campione, E., & Véronis, J. (2002). *A Large-Scale Multilingual Study of Silent Pause Duration*. 4.

Campione, E., & Véronis, J. (2005). Pauses and hesitations in French spontaneous speech. *DiSS*.

Candea, M. (2000). *Contribution à l'étude des pauses silencieuses et des phénomènes dits « d'hésitation » en français oral spontané. Etude sur un corpus de récits en classe de français.* [Phdthesis, Université de la Sorbonne nouvelle - Paris III]. https://tel.archives-ouvertes.fr/tel-00290143

Cuenca, M. H., & Barrio, M. M. (2010). Acoustic markers of prosodic boundaries in Spanish spontaneous alaryngeal speech. *Clinical Linguistics & Phonetics*, *24*(11), 859-869. https://doi.org/10.3109/02699206.2010.511404

Defossez G, et al. (2019). *Estimations nationales de l'incidence et de la mortalité par cancer en France métropolitaine entre 1990 et 2018. Volume 1 – Tumeurs solides. 1*, 372. Disponible à partir des URL : http://www.santepubliquefrance.fr/ ; https://geodes.santepubliquefrance.fr ; http://lesdonnees.e-cancer.fr/ ; https://www.e-cancer.fr/

Dejonckere, P., Bradley, P., Clemente, P. et al. A basic protocol for functional assessment of voice pathology, especially for investigating the efficacy of (phonosurgical) treatments and evaluating new assessment techniques . European Archives of Oto-Rhino-Laryngology 258, 77–82 (2001). https://doi.org/10.1007/s004050000299

Dodane, C., & Hirsch, F. (2018). L'organisation spatiale et temporelle de la pause en parole et en discours. *Langages*, *N° 211*(3), 5-12.

Dragičević, D., Jović, R. M., Kljajić, V., Vlaški, L., & Savović, S. (2019). Comparison of Voice Handicap Index in Patients with Esophageal and Tracheoesophageal Speech after Total Laryngectomy. *Folia Phoniatrica et Logopaedica*, 1-7. https://doi.org/10.1159/000502091

Duez, D. (2001). Signification des hesitations dans la production et la perception de la parole spontanée. *Revue parole*, *17*, 113-138.

Galant, C. (2019). *Contribution à l'étude des troubles du rythme de la parole après chirurgie pour cancer du larynx* [These de doctorat, Aix-Marseille]. http://www.theses.fr/s110991

Globlek, D., Stajner-Katusic, S., Musura, M., Horga, D., & Liker, M. (2004). Comparison of alaryngeal voice and speech. *Logopedics Phoniatrics Vocology*, *29*(2), 87-91. https://doi.org/10.1080/14015430410033218

Goldman, J.-P., François, T., Roekhaut, S., & Simon, A.-C. (2010). *Étude statistique de la durée pausale dans différents styles de parole*. XXVIIIèmes journées d'étude sur la parole (JEP 2010). https://dial.uclouvain.be/pr/boreal/object/boreal:81909

Graville, D. J., & Palmer, A. D. (2018). 17—Speech Pathology and Rehabilitation. In R. B. Bell, R. P. Fernandes, & P. E. Andersen (Éds.), *Oral, Head and Neck Oncology and Reconstructive Surgery* (p. 367-376). Elsevier. https://doi.org/10.1016/B978-0-323-26568-3.00017-8

Grosjean, F., & Deschamps, A. (1975). *Analyse contrastive des variables temporelles de l'anglais et du français : Vitesse de parole et variables composantes, phénomènes d'hésitation*. https://doi.org/10.1159/000259667

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

Harrus-Révidi, G. (2007). La laryngectomie, un cataclysme de l'oralité. *Champ psychosomatique*, *n° 48*(4), 145-152.

Hirsch, F., Didirková, I., Fauth, C., & Legou, T. (2018). Que cachent les pauses silencieuses en parole ? Une étude de cas. *Langages*, *N° 211*(3), 111-126.

Hurren, A., Miller, N., & Carding, P. (2019). Perceptual Assessment of Tracheoesophageal Voice Quality With the SToPS : The Development of a Reliable and Valid Tool. *Journal of Voice*, *33*(4), 465-472. https://doi.org/10.1016/j.jvoice.2017.12.006

Jacobson Barbara H., Johnson Alex, Grywalski Cynthia, Silbergleit Alice, Jacobson Gary, Benninger Michael S., & Newman Craig W. (1997). The Voice Handicap Index (VHI). *American Journal of Speech-Language Pathology*, *6*(3), 66-70. https://doi.org/10.1044/1058-0360.0603.66

Lalain, M., Legou, T., Fauth, C., Hirsch, F., & Didirkova, I. (2016, juillet). Que disent nos silences ? Apport des données acoustiques, articulatoires et physiologiques pour l'étude des pauses silencieuses. *JEP - Journées d'Etude sur la Parole*. https://hal.archives-ouvertes.fr/hal-01462180

Lalain, M., Mendonca-Alves, L., Espesser, R., Ghio, A., De Looze, C., & Reis, C. (2012). Lecture et Prosodie chez l'enfant dyslexique, le cas des pauses. *Journées d'Etudes sur la Parole*, 41-48. https://hal.archives-ouvertes.fr/hal-01510450

Laroche-Bouvy, D. (1984). Les pauses et les silences dans l'intéraction verbale. *Langage & société*, *29*(1), 27-37. https://doi.org/10.3406/lsoc.1984.1999

MacCallum, J. K., Cai, L., Zhou, L., Zhang, Y., & Jiang, J. J. (2009). Acoustic Analysis of Aperiodic Voice : Perturbation and Nonlinear Dynamic Properties in Esophageal Phonation. *Journal of Voice*, *23*(3), 283-290. https://doi.org/10.1016/j.jvoice.2007.10.004

Moerman, M.B.J., Martens, J.P., Van der Borgt, M.J. et al. Perceptual evaluation of substitution voices: development and evaluation of the (I)INFVo rating scale. Eur Arch Otorhinolaryngol 263, 183–187 (2006). https://doi.org/10.1007/s00405-005-0960-z

Moerman, M., Martens, J.-P., Crevier-Buchman, L., de Haan, E., Grand, S., Tessier, C., Woisard, V., & Dejonckere, P. (2006). The INFVo perceptual rating scale for substitution voicing : Development and reliability. *European Archives of Oto-Rhino-Laryngology and Head & Neck*, *263*(5), 435-439. https://doi.org/10.1007/s00405-005-1033-z

Moerman, M., Martens, JP. & Dejonckere, P. Application of the Voice Handicap Index in 45 patients with substitution voicing after total laryngectomy. Eur Arch Otorhinolaryngol 261, 423–428 (2004). https://doi.org/10.1007/s00405-003-0683-y

Moerman, M., Martens, J.-P., & Dejonckere, P. (2015). Multidimensional assessment of strongly irregular voices such as in substitution voicing and spasmodic dysphonia : A compilation of own research. *Logopedics Phoniatrics Vocology*, *40*(1), 24-29. https://doi.org/10.3109/14015439.2014.936497

Moerman, M., Pieters, G., Martens, JP. et al. Objective evaluation of the quality of substitution voices. Eur Arch Otorhinolaryngol 261, 541–547 (2004). https://doi.org/10.1007/s00405-003-0681-0

Most, T., Tobin, Y., & Mimran, R. C. (2000). Acoustic and perceptual characteristics of esophageal and tracheoesophageal speech production. *Journal of Communication Disorders*, *33*(2), 165-181. https://doi.org/10.1016/S0021-9924(99)00030-1

Motta, S., Galli, I., & Rienzo, L. D. (2001). Aerodynamic Findings in Esophageal Voice. *Archives of Otolaryngology–Head & Neck Surgery*, *127*(6), 700-704. https://doi.org/10.1001/archotol.127.6.700

Ng, M. L., Gilbert, H. R., & Lerman, J. W. (2001). Fundamental Frequency, Intensity, and Vowel Duration Characteristics Related to Perception of Cantonese Alaryngeal Speech. *Folia Phoniatrica et Logopaedica*, *53*(1), 36-47. https://doi.org/10.1159/000052652

Ng, Manwa L. (2011). Aerodynamic characteristics associated with oesophageal and tracheoesophageal speech of Cantonese. *International Journal of Speech-Language Pathology*, *13*(2), 137-144. https://doi.org/10.3109/17549507.2011.534176

Ng, Manwa L. (2019). The use of the Lombard Effect in Improving Alaryngeal Speech. *Journal of Voice*. https://doi.org/10.1016/j.jvoice.2019.07.007

Ordin, M. (2019). Speech rhythm as naturally occurring and culturally transmitted behavioral patterns. *Annals of the New York Academy of Sciences*, *1453*(1), 5-11. https://doi.org/10.1111/nyas.14234

PAROLE : Définition de PAROLE. (s. d.). Consulté 1 février 2021, à l'adresse https://www.cnrtl.fr/definition/parole

Ramanarayanan, V., Bresch, E., Byrd, D., Goldstein, L., & Narayanan, S. S. (2009). Analysis of pausing behavior in spontaneous speech using real-time magnetic resonance imaging of articulation. *The Journal of the Acoustical Society of America*, *126*(5), EL160-EL165. https://doi.org/10.1121/1.3213452

Skarnitzl, R., & Machač, P. (2009). Principles of Phonetic Segmentation. Phonetica, 68(3), 198-199. https://doi.org/10.1159/000331902

Smith, K. M., & Caplan, D. N. (2018). Communication impairment in Parkinson's disease : Impact of motor and cognitive symptoms on speech and language. *Brain and Language*, *185*, 38-46. https://doi.org/10.1016/j.bandl.2018.08.002

Stephen J. Cowley. (1994). Conversational functions of rhythmical patterning : A behavioural perspective. *Language & Communication*, *14*(4), 353-376. https://doi.org/10.1016/0271-5309(94)90026-4

Team, R. (2013). A language and environment for statistical computing. *Computing*, *1*. https://doi.org/10.1890/0012-9658(2002)083[3097:CFHIWS]2.0.CO;2

Teixeira, J. P., Fernandes, M. G., & Costa, R. A. (2012). Measure and Comparison of Speech Pause Duration in Subjects with Disfluency Speech. *Procedia Technology*, *5*, 812-819. https://doi.org/10.1016/j.protcy.2012.09.090

van Rossum, M. A. (2005). *Prosody in Alaryngeal Speech* [Thèse de doctorat, Utrecht University]. /paper/Prosody-in-Alaryngeal-Speech-Rossum/de9ce93718566a3d8b89c3d3c2cd20ae9348e3f3

van Rossum MA, de Krom G, Nooteboom SG, Quené H. "Pitch" accent in alaryngeal speech. J Speech Lang Hear Res. 2002 Dec;45(6):1106-18. https://doi.org/10.1044/1092-4388(2002/089)

van Rossum, M. A., Quené, H., & Nooteboom, S. G. (2008). Prosodic boundaries in alaryngeal speech. *Clinical Linguistics & Phonetics*, *22*(3), 215-231. https://doi.org/10.1080/02699200701847160

van Sluis, K.E., van der Molen, L., van Son, R.J.J.H. et al. Objective and subjective voice outcomes after total laryngectomy: a systematic review. Eur Arch Otorhinolaryngol 275, 11–26 (2018). https://doi.org/10.1007/s00405-017-4790-6

World Report on Disability 2011. (2011). World Health Organization. http://www.ncbi.nlm.nih.gov/books/NBK304079/

Yan, N., Lam, P., & Ng, M. (2012). Pitch Control in Esophageal and Tracheoesophageal Speech of Cantonese. *Folia phoniatrica et logopaedica : official organ of the International Association of Logopedics and Phoniatrics (IALP)*, *64*, 241-247. https://doi.org/10.1159/000342825

Yan, N., Ng, M. L., Wang, D., Zhang, L., Chan, V., & Ho, R. S. (2013). Nonlinear Dynamical Analysis of Laryngeal, Esophageal, and Tracheoesophageal Speech of Cantonese. *Journal of Voice*, *27*(1), 101-110. https://doi.org/10.1016/j.jvoice.2012.06.009

Yunusova, Y., Graham, N. L., Shellikeri, S., Phuong, K., Kulkarni, M., Rochon, E., Tang-Wai, D. F., Chow, T. W., Black, S. E., Zinman, L. H., & Green, J. R. (2016). Profiling Speech and Pausing in Amyotrophic Lateral Sclerosis (ALS) and Frontotemporal Dementia (FTD). *PLoS ONE*, *11*(1). https://doi.org/10.1371/journal.pone.0147573

Zenga, J., Goldsmith, T., Bunting, G., & Deschler, D. G. (2018). State of the art : Rehabilitation of speech and swallowing after total laryngectomy. *Oral Oncology*, *86*, 38-47. https://doi.org/10.1016/j.oraloncology.2018.08.023

Annexes

Annexe 1 : Sujets opérés d'une laryngectomie totale

Sujets	Sexe	Age	Classification	Date LT	Distance à LT (en mois)	CT	RT	Filtre
1	M	72	T4 N0 M0	2011	48	oui	oui	non
2	M	72	pT4a pN1	2016	6	oui	oui	oui
3	M	67	NC	2008	84	oui	oui	oui
4	M	74	T3 N0 M0	2012	36	oui	non	non
5	M	63	cT4a N0 M0	2014	12	oui	oui pré-op	oui
6	M	73	T4a N0 M0	2003	156	oui	oui	oui
7	M	56	pT4a N0	2016	6	oui	oui	oui
8	M	77	NC	1994	252	non	oui	non
9	F	42	T4 N0 M0	2011	48	oui	oui	oui
10	M	73	T3 N0 M0	2014	18	oui	oui pré-op	oui
11	M	53	T4 N0 Mx	2008	84	oui	oui pré-op	non
12	M	80	T3 N0 Mx	2010	60	oui	oui	non
13	M	65	NC	2006	108	oui	non	oui
14	M	83	T4 N0 M0	2014	12	oui	oui pré-op	non
15	M	74	T3 N1 M0	2007	96	oui	oui pré-op	non
16	M	71	T4 N0 M0	2008	84	oui	oui	oui
17	M	73	T3 N1 M0	2012	36	oui	oui	non
18	M	60	NC	2006	108	oui	oui	oui
19	M	71	T3 N0 M0	2010	60	oui	oui	non
20	M	50	T4 N1 M0	2013	24	oui	oui	non
Moyenne		72			66,9			
Médian		72			54			
Min		67			6			
Max		74			252			
Ecart-type		63			59,692104			

Annexe 2 : Sujet contrôles

SUJET	SEXE	AGE
1	M	62
2	M	61
3	M	65
4	M	79
5	M	78
6	M	66
7	M	68
8	M	58
9	M	72
10	M	62
11	M	66
12	F	67
13	M	59
14	M	57
15	M	70
16	M	68
17	M	70
18	M	78
19	M	59
20	M	82
21	M	64
Moyenne		67,1904762
Médiane		66
Min		57
Max		82
Ecart-type		7.12219

Annexe 3a : Analyse qualitative des échantillons de parole semi-spontanée – Sujets LT

De manière générale, on observe que les sujets laryngectomisés ont tendance à coupler des fonctions au sein d'une même pause afin d'en optimiser l'occurrence (pblocsy, pbloches, pblocac, etc). Ces pauses hybrides semblent d'ailleurs plus longues que les simples, particulièrement celles avec une visée syntaxique. Ces dernières semblent placées selon l'organisation syntaxique et sémantique de la phrase. D'ailleurs, les pauses de blocage syntaxiques situées entre deux phrases semblent souvent plus longues, peut-être car elles englobent en réalité également des fonctions de respiration et de planification.

Les patients n'ayant pas une bonne maîtrise de leur voix ont tendance à lister les informations plutôt qu'à faire des phrases. De même, ils ont tendance à syllaber et, de fait, à produire un plus grand nombre de pauses. On note même deux patients présentant des pauses intra-syllabiques, mais cela reste rare. Nous remarquons un certain nombre de patients présentant un souffle trachéal important, impactant la qualité de la parole ainsi que sa perception.

Toutefois, certains patients sont capables de produire des variations de hauteurs et des intonations. D'autres encore présentent des pauses d'hésitation remplies, ce qui témoigne d'une bonne maîtrise vocale. Enfin, on observe un allongement des syllabes situées aux frontières syntaxiques (devant des psy, pblocsy, prespisy ou encore tdp), mais également parfois devant des pauses à visée cognitive (phes, phesr ou pac). Cela montre une envie de coller aux schémas de parole retrouvés chez les sujets sains – dont l'hésitation est marquée par un allongement syllabique afin de conserver le tour de parole – mais aussi l'allongement prosodique final typique du français.

SUJET	GROUPE	SEXE	COMMENTAIRES
1	LT	M	Ce patient contrôle très bien son souffle et semble bien gérer les blocages. Il est capable de produire plusieurs mots avec un même blocage, et lorsqu'il est obligé d'en faire au milieu de la phrase, il arrive à les rendre discrets. Il possède une bonne prosodie, avec des syllfinale allongée comme chez les sujets sains pour marquer les fins de phrases.
2	LT	M	Présence d'un fort souffle trachéal. Les blocages se font sur l'inspiration et sont couverts par son fort bruit. Il syllabe les mots longs, et un blocage permet la production de 2 mots maximum, dont la fin est généralement très soufflée. Toutefois, il produit des phrases et ne liste pas. On note également de longues pauses d'hésitation : cognitives ou mécaniques ?
3	LT	M	Le patient bloque en inspirant. Il présente une voix plutôt fluide même si elle reste robotique. On note un souffle trachéal important et des pauses parfois très longues : reprise inspiratoire et temps de préparation au blocage et à l'articulation.
4	LT	M	Le patient a une parole globalement fluide. Il lui arrive de produire des pauses respiratoires d'hésitation assez longues, indiquant une

			reprise inspiratoire longue et une planification coocurrente. Quand il prend des pauses trop courtes il se retrouve rapidement à bout de souffle. Peut parfois faire quelques variations de hauteur de fréquence à la fin des phrases. Utilise efficacement ses pauses syntaxiques pour bloquer. Blocages sonores.
5	LT	M	Ce patient présente une voix de mauvaise qualité, avec un fort souffle trachéal. Il produit ses blocages en inspirant. Il tente de respecter les pauses syntaxiques en les allongeant par rapport à celles purement physiologiques.
6	LT	M	Bonne maîtrise du souffle et de la phonation. Il a tendance à avoir une prosodie robotique, excepté en fin de phrase où il produit une intonation descendante. Production très succincte, il va à l'essentiel, sensation de liste même s'il produit des phrases. Il produit quelques « euh » d'hésitation, ce qui montre une maîtrise des pauses et des blocages.
7	LT	M	Patient avec un voix moyennement maîtrisée et un souffle trachéal prégnant. On note des pauses d'injection sur certaines consonnes, et de longues pauses d'hésitation (planification cognitive, motrice ?). Production très succincte mais informative et efficace.
8	LT	M	Pas de souffle trachéal systématique, mais on entend un souffle en fin de phrase (à bout de souffle ?). Blocages très audibles, avec parfois nécessité d'en faire plusieurs avant d'être capable de produire son énoncé. Beaucoup de reprises inspiratoires, sûrement nécessaires aux blocages. De fait, il a tendance à lister. On note que les syllfinales semblent allongées, mais pas d'intonation descendante. De fait, la parole n'est pas fluide et on sent les efforts que lui coûte la phonation.
9	LT	F	Patiente avec une voix robotique mais présentant une bonne fluence. Elle peut produire des intonations et des hésitations très naturelles. La syntaxe est respectée et les pauses à cet effet sont marquées. Elle produit des blocages mais aussi des injections (avec la consonne k principalement), qui semblent plus courtes que les premières.
10	LT	M	Voix très mauvaise, fort souffle trachéal, qui engendre l'adjonction d'une grande quantité de pauses. Cette mauvaise maîtrise l'oblige à syllaber et à lister (peu de syntaxe et de prosodie). La parole est désonorisée, les blocages très difficiles à produire, ce qui donnent des pauses parfois extrêmement longues (jusqu'à 4 secondes). On observe des pauses intra-syllabiques.
11	LT	M	Très bonne voix. Ce patient procède à ses blocages en s'appuyant sur une inspiration (qui est audible). Toutefois on n'entend pas de souffle trachéal lors de la phonation. Il produit des phrases longues et complexes. Un blocage lui permet la production de plusieurs mots, sa parole est donc relativement fluide. Il possède une bonne prosodie, avec des syllabes finales à intonation descendante. On observe de nombreux « euh » d'hésitation (phesr) qui témoignent d'une maîtrise de la voix oro-œsophagienne.
12	LT	M	On constate un souffle trachéal très présent. Des blocages sont nécessaires tous les deux mots au minimum, et ceux-ci entraînent des pauses plus longues que chez d'autres parleurs. Les pblocsy sont de fait encore plus longues (jusqu'à plus de 5 secondes). Cela semble principalement dû à une planification mécanique plutôt que cognitive (difficulté à bloquer et réservoir d'air faible).

			On note toutefois une intonation descendante et un allongement de la syllfinale en fin de phrase. De plus, malgré les nombreuses pauses ajoutées au discours et l'effort mis en jeu, le patient fait des phrases et ne liste pas.
13	LT	M	Patient produisant ses blocages sur inspiration, il présente une voix de qualité moyenne mais fluide. On a l'impression qu'il force pour ne pas avoir à rebloquer en fin de mot/phrase. On remarque quelques intonations/modifications de hauteur, ainsi que des pauses d'hésitation remplies. Les blocages sont nombreux (quasiment après chaque mot) et très sonorisés. On note également un léger souffle trachéal. Enfin, il produit quelques injections sur la consonne p.
14	LT	M	Patient en récidive lors de l'enregistrement : on note de fait une mauvaise voix, avec des blocages très audibles, ne permettant que peu de production. Le souffle trachéal est très présent lors de la production des phonèmes, avec presque une sensation d'asphyxie. S'il arrive parfois à produire un mot entier avec un blocage, il est quasi systématiquement obligé de faire une pause de blocage après chaque syllabe (pause intra-mots). Les pauses sont longues (planification cognitive et mécanique ?). Il est très hypospontané, l'orthophoniste intervient après presque chaque phrase pour relancer et demander des précisions. Sa description est très peu détaillée, sûrement dû à l'effort que lui coûte le fait de parler.
15	LT	M	Le souffle n'est pas dissocié de la production phonatoire, on note donc un souffle trachéal audible lorsqu'il parle. Toutefois, la gestion des blocages semble efficace, avec la possibilité de produire plusieurs mots avec une prise d'air dans l'œsophage. Son discours est fluide, avec un débit rapide, très peu impacté par les pauses physiologiques produites. Celles-ci sont d'ailleurs relativement courtes, le blocage est direct et efficace. Ils sont intelligemment positionnés, ce qui donne une perception agréable et fluide de la parole. On constate une prosodie, avec des intonations (syllfinales allongées et descendantes). Toutefois, on sent un effet de longueur, avec une baisse d'intensité sur les fins de phrases – sur laquelle le souffle trachéal devient plus audible – mais aussi sur la fin de l'enregistrement, où la qualité de la voix semble moins bonne qu'au début.
16	LT	M	Présence d'un souffle trachéal important. La gestion du souffle n'étant pas efficace, il procède parfois à de longues pauses de reprise inspiratoire. Le blocage est parfois très discret. La voix est robotique mais fluide, et il arrive à produire des intonations et des hésitations naturelles. La syntaxe est assez élaborée et on repère facilement les pauses à cet effet (plus longues ?).
17	LT	M	Patient présentant lui aussi une récidive de son cancer. Le souffle trachéal est très présent, sa voix oro-œsophagienne est de fait couplée avec ce souffle. Il a besoin de pbloc régulières, à peu près tous les 2 mots. Il est parfois capable de produire plus de mots en un blocage, mais cela induit ensuite une pause de blocage plus longue, avec un besoin de reprise inspiratoire pour produire un autre blocage. Il produit toutefois des phrases sans lister, et sa description est courte mais précise. Malgré une mauvaise qualité de voix, il produit un abaissement de la fréquence sur les fins de phrase.

18	LT	M	Voix mauvaise, avec souffle trachéal et mauvaise gestion de la sonorisation. Les pauses sont longues (surtout inter-phrases) et très nombreuses : on note des pauses intra-mots, voire intra-syllabiques. Les blocages sont très audibles, et parfois multiples. Tout cela entraîne une syllabation, et donc des phrases très courtes avec peu de syntaxe et d'informations (l'orthophoniste relance régulièrement). Enfin, il produit des injections sur la consonne p devant certains phonèmes (les fricatives principalement).
19	LT	M	On observe un souffle trachéal important, de multiples blocages en une pause et des bruits de bouche, entraînant des bruits parasites. La voix est mauvaise, presque aucune intonation. Il liste les informations, sans parfois faire de phrases. Sa mauvaise gestion des blocages entraîne donc le besoin d'en faire plusieurs avant d'arriver à sonoriser, mais aussi l'occurrence de pauses intra-mots.
20	LT	F	Bonne voix. On observe beaucoup de « euh » d'hésitation, indicateurs d'une qualité et d'une aisance de production. On n'entend pas de souffle trachéal, et les blocages permettent la production de plusieurs mots, ce qui donne une parole plutôt fluide. Les blocages sont bien positionnés en intra-phrases, mais on note des pblocsy inter-phrases plutôt longues (peut-être utilisées également pour une planification cognitive). Il semble qu'elle essaie de « rentabiliser » chaque pause pour y assurer les fonctions physiologiques, syntaxiques et cognitives en simultané. Les syllfinales sont allongées et la fréquence descend pour signifier la fin de phrase. On constate également une prosodie qui amplifie la sensation de fluidité.

Annexe 3b : Analyse qualitative des échantillons de parole semi-spontanée – Sujets sains

Comparés aux sujets laryngectomisés, ils semblent produire moins de pauses, des phrases plus longues et plus complexes, et leurs descriptions sont dans l'ensemble plus précises (peut-être car ils ne sont pas restreints par leur parole). D'ailleurs, les sujets utilisant un langage « soutenu » procèdent à davantage de pauses à visée syntaxique. On note plus de pauses d'hésitations remplies que les sujets LT, ainsi que des allongements de syllabe pour marquer cette hésitation. D'ailleurs, ces « phesr » sont observables chez quasiment tous les patients contrôles, tandis que seuls quelques sujets alaryngés en sont capables.

On observe que les pauses sont généralement utilisées de façon à être optimisées par les sujets sains, si bien qu'on peut trouver les 3 fonctions principales (physiologique, syntaxique et cognitive) dans une même pause. De manière générale, ils profitent des pauses syntaxiques pour procéder aux reprises inspiratoires.

SUJET	GROUPE	SEXE	COMMENTAIRES
21	Sain	M	Ce sujet produit beaucoup de pauses d'hésitation, aussi bien vides que pleines. De plus, ces hésitations se retrouvent aussi dans l'allongement de certains phonèmes finaux.
22	Sain	M	Il ne finit pas toujours ses phrases ce qui provoque quelques pauses d'autocorrection. Ce sujet produit une pause respiratoire syntaxique très longue (plus de 3 secondes) et un nombre important de pauses d'hésitations remplies. On remarque des commentaires de type "je crois, peut-être, en fait"
23	Sain	M	Ses pauses syntaxiques sont toutes accompagnées d'une reprise inspiratoire. Il y a quelques pauses d'hésitation et bien sûr un allongement des syllabes finales.
24	Sain	M	Débit normal mais interrompu très souvent par des pauses d'hésitations remplies. En voulant être détaillé il perd un peu d'efficacité dans sa description.
25	Sain	M	Débit rapide, phrases très longues et complexes, avec beaucoup de détails. Principalement des pauses syntaxiques avec reprise inspiratoire. Celles d'hésitation sont remplies, les vides sont très rares.
26	Sain	M	Ce sujet procède à beaucoup d'allongement syllabique pour marquer son hésitation. Ses inspirations donnent une impression de « schlintement ».
27	Sain	M	Discours lent et posé, avec quelques pauses de focalisation semblant mettre en avant les éléments qu'il trouve importants dans sa description. Beaucoup de pauses cognitives (phes et phesr) servant à la planification de la parole.
28	Sain	M	Ce sujet produit d'importants allongements de syllabes et de pauses d'hésitations remplies indiquant une planification assez conséquente.
29	Sain	M	On a l'impression qu'il dicte sa description. Il produit des pauses qui ne semblent pas être à visée syntaxique mais qui séparent les groupes sémantiques. On note beaucoup de « phes » de planification.

30	Sain	M	Discours fluide mais tendance à énumérer des groupes nominaux pour la description, moins de phrases de fait. Toutefois, cette énumération est mise en exergue par l'intonation et la prosodie. Pas mal de pauses d'hésitation (vides ou remplies) qui montre le processus de planification cognitive.
31	Sain	M	Beaucoup de pauses d'hésitation (principalement remplies). En revanche, moins de pauses syntaxiques, il enchaîne les phrases et reprend plutôt son inspiration en intra-phrase. De fait, la prosodie est essentielle pour délimiter les phrases (allongement de la syllabe finale). La description est très précise.
32	Sain	F	Le sujet a un débit important mais reste clair et efficace. Informations très détaillées mais cohérentes et compréhensibles. Elle produit des phrases assez longues et avec une bonne syntaxe.
33	Sain	M	Les reprises inspiratoires sont très espacées, avec beaucoup de parole entre. Les phrases sont principalement délimitées par la prosodie, avec l'allongement de la dernière syllabe et l'abaissement de la fréquence simultanée. Description concise mais précise. On note des pauses d'hésitations.
34	Sain	M	Ce sujet produit des allongements syllabiques finaux lors d'énumération, ce qui fait qu'on n'observe pas de pauses syntaxiques derrière. On note aussi une pause de focalisation isolée.
35	Sain	M	Les reprises inspiratoires se font sur les pauses syntaxiques. On note quelques « euh » d'hésitation (phesr). Du fait de ses phrases plus longues et complexes, on obtient une description plus courte mais plus précise qu'un sujet LT ayant produit plus de phrases.
36	Sain	M	Ce sujet semble narrer sa description, avec des pauses interphrases (syntaxiques et respiratoire) longues, des allongements de syllabes importants. Reprises inspiratoires bien audibles avec même une aspration de salive concurrente.
37	Sain	M	Une pause d'hésitation remplie commence presque toutes ses phrases, ou du moins on retrouve une hésitation dans les premiers mots de chaque phrase.
38	Sain	M	Ce patient énumère les éléments de l'image plus qu'il ne la décrit. De fait, on observe une liste d'adjectif et beaucoup de pauses syntaxiques et d'hésitation les démarquant. Ses pauses sont longues. Le débit est donc très lent, avec une prosodie assez monotone.
39	Sain	M	Les pauses syntaxiques servent à chaque fois à la reprise inspiratoire. On ne retrouve pas de respiration ailleurs. Quelques pauses d'hésitation.
40	Sain	M	Ses pauses respiratoires sont souvent optimisées au moment des pauses syntaxiques. On note de nombreuses pauses de planification (hésitation, autocorrection, etc).
41	Sain	M	Beaucoup d'hésitation, qui engendre soit des pauses silencieuses, généralement courtes, soit des allongements de phonèmes : isolés (« euh ») ou prolongeant un mot (syllfinale).

Annexe 4 : exemple d'images support de la description.

Résumé

Les patients utilisant la voix oro-œsophagienne (LT) se plaignent de leur voix, mais aussi de la fluidité de leur parole. Cette dernière a été très peu décrite, et l'impact des pauses n'a été mis en évidence qu'en termes de quantité, plus importante. Camille Galant a démontré en lecture que l'ajout de pauses par le sujet laryngectomisé perturbait le rythme de leur parole, mais qu'ils les adaptaient pour respecter sa structure syntaxique.

Notre mémoire étudie l'organisation temporelle de la parole oro-œsophagienne en parole semi-spontanée, et plus précisément la distribution des pauses. Nous émettons deux hypothèses : les sujets laryngectomisés produiraient plus de pauses que les sujets sains, mais ces dernières auraient une durée différente selon leur fonction dans la parole. 20 sujets laryngectomisés ont été enregistrés et comparés à 21 sujets sains.

Nous avons mis en évidence que les sujets LT produisent plus de pauses que les sains, mais les optimisent (durée, quantité, cooccurrence de fonctions). On observe un respect de la structure syntaxique du discours. La situation de description d'image révèle plus de types de pauses qu'en lecture. On note une durée plus conséquente des pauses « hybrides », support de différentes fonctions, par rapport aux pauses « simples ». Le sujet laryngectomisé fait donc bien varier la durée de ses pauses selon leur fonction. Le débit articulatoire est plus lent pour les LT mais la vitesse articulatoire est équivalente chez les deux groupes.

Nos hypothèses sont validées par nos descriptions phonétiques de l'organisation rythmique de la parole semi-spontanée, mais notre travail comporte différentes limites qui pourront être exploitées afin d'appuyer leur véracité. L'étude rythmique de la parole oro-œsophagienne n'en est qu'à ses débuts, des connaissances approfondies pourraient permettre une évaluation et une rééducation adaptées.

Mots clés : *laryngectomie totale, parole oro-œsophagienne, rythme, pause, orthophonie*

Abstract

Patients using the oroesophageal (LT) voice complain about their voice, but also about the fluidity of their speech. This has been poorly described, and the impact of the pauses was only highlighted in terms of the larger quantity. Camille Galant demonstrated in reading that the addition of pauses by the laryngectomized subjects disrupted the rhythm of their speech, but that they adapted them to respect its syntactic structure.

Our thesis studies the temporal organization of oroesophageal speech in semi-spontaneous speech, and more precisely the distribution of pauses. We put forward two hypotheses: laryngectomized subjects would produce more pauses than healthy subjects, but the latter would have a different duration depending on their function in speech. 20 laryngectomized subjects were recorded and compared to 21 healthy subjects.

We have demonstrated that LT subjects produce more pauses than healthy subjects, but optimize them (duration, quantity, co-occurrence of functions). We observe a respect for the syntactic structure of speech. The picture description situation reveals more types of pauses than in reading. We note a longer duration of "hybrid" pauses, support for different functions, compared to "simple" pauses. The laryngectomized subject therefore varies the duration of his breaks according to their function. Articulatory flow is slower for LT but articulatory speed is equivalent in both groups.

Our hypotheses are validated by our phonetic descriptions of the rhythmic organization of semi-spontaneous speech, but our work has different limitations that can be exploited to support their veracity. The rhythmic study of oroesophageal speech is only in its infancy, in-depth knowledge could allow an appropriate assessment and rehabilitation.

Keywords: *total laryngectomy, orœsophageal speech, rhythm, pause, speech therapy*