

Étude des liens acoustico-moteurs après cancer oral ou oropharyngé, via la réalisation d'un inventaire phonémique automatique des consonnes

Mathieu Balaguer^{1,2} Lucile Gelin^{1,3} Clémence Devoucoux² Camille Galant⁴ Muriel Lalain⁴
Alain Ghio⁴ Jérôme Farinas¹ Julien Pinquier¹ Virginie Woisard^{2,5}

(1) IRIT, Université de Toulouse, CNRS, Toulouse INP, UT3, Toulouse, France

(2) Hôpital Larrey, Toulouse, France

(3) Lalilo, Paris, France

(4) Aix-Marseille Université, CNRS, LPL, UMR 7309, Aix-en-Provence, France

(5) Laboratoire de NeuroPsychoLinguistique, UR 4156, Université de Toulouse, Toulouse, France

mathieu.balaguer@irit.fr

RESUME

En cancérologie ORL, le lien entre anatomie et déficit de parole est étroit en raison de l'impact de la pathologie et de son traitement sur les structures anatomiques en jeu dans la production de parole. Pourtant, les corrélations entre scores moteurs et évaluation perceptive restent faibles. L'utilisation de systèmes automatiques dédiés à la reconnaissance de phonèmes pourrait permettre d'obtenir de nouveaux résultats. L'objectif est d'étudier les liens entre scores moteurs et production phonémique via un système de reconnaissance automatique de phonèmes appliqué à une tâche de production de pseudo-mots. Après réalisation d'un inventaire phonémique par sujet, le taux d'occlusives reconnues est significativement plus faible en cas d'atteinte des structures. Certains mécanismes de compensation ont également pu être mis en évidence, notamment au niveau de la production de consonnes labiodentales, plus élevée en cas d'atteinte de la langue ou de la mâchoire.

ABSTRACT

Analysis of acoustico-motor links after oral or oropharyngeal cancer, using an automatic phonemic inventory of consonants

In ENT oncology, the link between anatomy and speech deficit is close, due to the impact of the pathology and its treatment on the anatomical structures involved in speech production. However, correlations between motor scores and perceptual evaluation remain weak. The use of automatic systems dedicated to phoneme recognition could lead to new results. The objective is to analyze the links between motor scores and phonemic production using an automatic phoneme recognition system applied to a pseudoword production task. After completing a phonemic inventory for each subject, the rate of occlusives recognized was significantly lower in cases of structural impairment. Compensatory mechanisms were also demonstrated, notably in the production of labiodental consonants, which was higher in cases of tongue or jaw damage.

MOTS-CLES : trouble de la parole, cancérologie, évaluation automatique, phonèmes, capacités motrices

KEYWORDS : speech disorder, oncology, automatic assessment, phonemes, motor skills

1 Contexte

Les cancers de la cavité buccale ou de l'oropharynx sont des cancers fréquents ([Lapôte-ledoux et al., 2023](#)). La question de la vie après cancer est essentielle, et il convient de s'intéresser aux fonctions des voies aérodigestives supérieures (VADS) particulièrement dégradées en raison de la localisation de ces cancers. La fonction de parole devient ainsi un sujet crucial en termes de réhabilitation, à cause de l'impact fonctionnel et psychosocial qu'un trouble de parole va avoir sur les personnes ([Mlynarek et al., 2008](#); [Reich, 2009](#)).

Les évaluations cliniques de la parole sont majoritairement menées par les orthophonistes ([Pommée et al., 2022](#)). En cancérologie des VADS, peu d'outils d'évaluation existent à l'heure actuelle ([Ghio et al., 2016](#)), alors qu'il s'agit pourtant du symptôme le plus fréquent ([Plisson et al., 2017](#)). L'évaluation clinique de la parole comprend habituellement deux volets. La partie analytique consiste en un bilan des structures anatomiques en vue de mettre en évidence un déficit d'amplitude, de tonus, de force motrice ou de sensibilité. La partie fonctionnelle concerne l'évaluation perceptive pour caractériser, sur des tâches de production de parole, les répercussions des déficits moteurs ou sensitifs sur les productions orales ([Middag, 2013](#)). Ici, l'intelligibilité est particulièrement ciblée dans les évaluations. Elle est définie comme la capacité à « *reconstruire un énoncé au niveau acoustico-phonétique* » ([Pommée et al., 2021](#)), en d'autres termes à décoder un signal de parole en éléments phonémiques sans mise en œuvre des mécanismes cognitifs de restauration du message par l'auditeur ([Ghio et al., 2018](#)).

Ainsi, même si en cancérologie le lien entre anatomie et déficit fonctionnel sur la parole est étroit, la corrélation reste faible entre les scores fonctionnels d'intelligibilité attribués perceptivement et les scores analytiques issus du bilan moteur ([Lazarus et al., 2013](#)). Les avancées récentes dans le champ de l'analyse automatique de la parole peuvent désormais permettre d'envisager une application clinique de ces techniques, en palliant certaines caractéristiques inhérentes à l'évaluation perceptive (notamment la variabilité inter et intra-juges ([Fex, 1992](#); [Middag et al., 2008](#))). Des études ont par exemple déjà montré l'intérêt d'une mesure de performance de systèmes de reconnaissance automatique de parole, comme mesure d'intelligibilité ([Christensen et al., 2012](#); [Doyle et al., 1997](#); [Maier et al., 2010](#)). L'utilisation de systèmes de reconnaissance automatique de phonèmes pour dresser un inventaire des phonèmes reconnus peut permettre ainsi d'ouvrir de nouvelles perspectives dans l'étude du lien entre parole et scores moteurs.

L'objectif est d'étudier les liens entre scores moteurs (cavité buccale, oropharynx) et production phonémique (via la constitution d'un inventaire phonémique des consonnes par utilisation de système de reconnaissance automatique de phonèmes) après traitement d'un cancer de la cavité buccale ou de l'oropharynx.

2 Matériel et méthodes

2.1 Schéma d'étude

Cette étude prospective observationnelle s'inscrit dans le cadre du projet PHRIP DAPADAF-E (PHRIP-19-0004). Toutes les procédures ont été effectuées conformément à la déclaration d'Helsinki de 1964 et à ses amendements ou à des normes éthiques comparables. Chaque sujet a été informé à l'avance de l'objectif de l'étude et a reçu une fiche d'information. Les sujets ont confirmé leur non-opposition à la collecte de données et l'utilisation de la recherche dans le projet DAPADAF-E. Les fichiers audio enregistrés ont été récupérés via le GIS Parolothèque, de même que les données individuelles, cliniques et de traitement.

2.2 Population

Les patients venant en consultation ORL ou dans un service de soins de suite et de réadaptation ORL entre novembre 2021 et décembre 2023, sur les sites d'inclusion des Hôpitaux de Toulouse ou de l'Assistance Publique – Hôpitaux de Marseille, ont été invités à participer à cette étude.

Les critères d'inclusion étaient : être majeurs, francophones natifs, avoir été traités pour un cancer de la cavité buccale ou de l'oropharynx (chirurgie et/ou radiothérapie et/ou chimiothérapie) et être en rémission clinique depuis au moins 6 mois (caractère chronique et stable des troubles). N'ont pas été inclus les patients dont la fatigabilité ne permet pas la passation des épreuves, ou présentant une pathologie associée potentiellement responsable de trouble de parole (bégaiement, trouble neurologique...).

2.3 Corpus

2.3.1 Scores analytiques moteurs

En raison de l'absence de test exhaustif et validé en cancérologie ORL, le protocole DAPADAF-E s'appuie sur le bilan clinique issu de la BECD (Batterie d'Évaluation Clinique de la Dysarthrie, [\(Auzou & Rolland-Monnoury, 2019\)](#)), et particulièrement sur l'épreuve d'examen moteur.

Ainsi, pour chaque structure anatomique d'intérêt (lèvres, joues, mâchoire, langue et vélopharynx), deux scores sont obtenus : un score de synthèse analytique (correspondant aux épreuves du domaine non-verbal : tâches motrices hors production de parole) et un score de synthèse fonctionnel (domaine verbal : tâches associées à une production de parole). Chacun de ces scores est coté sur le principe d'une échelle de Likert à 5 niveaux : 0 correspondant à l'absence d'anomalie, 1 à une anomalie discrète ou rare, 2 à une anomalie modérée ou occasionnelle, 3 à une anomalie marquée ou fréquente et 4 à une anomalie sévère ou quasi permanente. Pour les besoins de cette étude et pour obtenir un nombre plus important de sujets dans chaque classe, les scores moteurs (analytiques et fonctionnels) ont été dichotomisés, de façon à obtenir deux valeurs pour chacun : « *pas d'anomalie, anomalie discrète ou modérée* » (0 : scores entre 0 et 2) et « *anomalie marquée, sévère ou quasi permanente* » (1 : scores 3 ou 4).

2.3.2 Enregistrements de parole

Les sujets ont été enregistrés sur une tâche de production de pseudo-mots respectant les règles phonotactiques du français (par exemple : *crerquin, ruflu...*). Ils étaient installés devant un écran affichant le pseudo-mot à produire et portaient un casque dans lequel une voix de synthèse prononçait le pseudo-mot, de façon à ce qu'ils aient accès à la double modalité lecture et répétition pour produire le pseudo-mot cible. Chaque sujet a été enregistré sur deux listes de 52 pseudo-mots, différentes entre elles et pour chaque patient. Ces listes ont été mises au point par l'équipe du Laboratoire Parole et Langage d'Aix-en-Provence ([Ghio et al., 2022, 2018](#)).

Les enregistrements ont été réalisés au moyen d'un micro Neumann TLM 102, protégé par une bonnette et un filtre anti pop, et connecté à une interface audio RME Fireface UC. L'interface était ensuite reliée à un ordinateur portable sur lequel le logiciel LiveIntel (également mis au point par le LPL) était installé et qui gérait l'enregistrement de la tâche. Les sujets étaient enregistrés dans une

salle de consultation calme. Les enregistrements étaient au format 48 kHz 16 bits, puis ils ont été rééchantillonnés en 16 kHz car il s'agit de l'usage dans le traitement automatique.

2.4 Analyse automatique de la parole

Chaque fichier, comprenant la production de 52 pseudo-mots, a été ensuite segmenté par le détecteur d'activité vocale WebRTC-VAD¹, afin d'obtenir des segments de parole courts (de durée inférieure à 25 secondes) pouvant être gérés par le système de reconnaissance de phonèmes.

Les fichiers segmentés ont ensuite été donnés à un système de reconnaissance automatique de phonèmes de type Transformer CTC, développé par les auteurs de cette étude et entraîné sur le corpus francophone CommonVoice, ici constitué de 148,9 heures d'enregistrements de parole lue par 1 276 locuteurs (avec 420 secondes d'enregistrements par locuteur en moyenne). Le décodage a abouti à la reconnaissance de la séquence des phonèmes produits, parmi 33 possibilités pour chaque phonème sur lequel le Transformer CTC a été entraîné :

- 18 consonnes : /p/, /t/, /k/, /b/, /d/, /g/, /f/, /s/, /ʃ/, /v/, /z/, /ʒ/, /m/, /n/, /ɲ/, /ŋ/, /l/, /ʀ/;
- 12 voyelles, dont /ã/, /ɛ/, /e/, /o/, /ɔ/, /i/, /ĩ/, /y/, /u/, et les archiphonèmes /A/ (/a/, /ɑ/), /E/ (/ə/, /œ/, /ø/) et /Ê/ (/ē/, /œ̃/);
- et trois semi-consonnes : /j/, /w/, /ɥ/.

Puis, sur l'ensemble des deux listes produites par sujet (deux fois 52 soit 104 pseudo-mots par sujet), un inventaire phonémique ciblé sur les consonnes a été dressé afin d'obtenir le nombre total d'occurrences pour chaque phonème.

Enfin, dix indicateurs ont été calculés suite à cet inventaire phonémique, concernant les taux de :

- Consonnes reconnues (*csn*) : nombre total de consonnes / nombre total de phonèmes ;
- Occlusives reconnues (*occ*) : nombre total d'occlusives / nombre total de consonnes ;
- Bilabiales (*bilabial*) : nombre total de consonnes bilabiales / nombre total de consonnes ;
- Labiodentales (*labiodent*) : nombre total de consonnes labiodentales / nombre total de consonnes ;
- Alvéolaires (*bilabial*) : nombre total de consonnes alvéolaires / nombre total de consonnes ;
- Post-alvéolaires (*postalv*) : nombre total de consonnes post-alvéolaires / nombre total de consonnes ;
- Palatales (*palatal*) : nombre total de consonnes palatales / nombre total de consonnes ;
- Vélares (*velaire*) : nombre total de consonnes vélares / nombre total de consonnes ;
- Sourdes (*sourde*) : nombre total de consonnes sourdes / nombre total de consonnes ;
- Consonnes orales (*corale*) : nombre total de consonnes orales / nombre total de consonnes.

2.5 Analyses statistiques

Les analyses statistiques ont été réalisées au moyen du logiciel Stata 16.1 (StataCorp. 2019. Stata statistical software: release 16. College Station, TX: StataCorp LLC). Un seuil de significativité à 5 % a été choisi pour l'ensemble des analyses. En raison du caractère non-paramétrique des données, le test de Mann-Whitney a été réalisé pour tester l'absence de différence significative de proportion des différentes catégories phonémiques selon le niveau d'anomalie motrice.

¹ Site : <https://github.com/wiseman/py-webrtcvad>, consulté le 02/02/2024

3 Résultats

3.1 Description de la population

Soixante-dix-sept sujets ont été inclus dans cette étude, majoritairement de sexe masculin (50/77, 65 %). La proportion de sujets avec tumeurs de petit et de grand volume est équilibrée (T1+T2 = 32/64, 50 % ; T3+T4 = 32/64, 50 % ; 13 données manquantes). La majorité des localisations tumorales se situent dans la cavité buccale (langue : 38, plancher buccal : 18, bouche : 14, gencives : 5 et glandes salivaires : 1) avec certains patients présentant plusieurs sites tumoraux. 30 sujets ont présenté un cancer de l'oropharynx. 63 sujets ont été traités par chirurgie (82 %), 66 (86 %) par radiothérapie et 41 (53 %) par chimiothérapie.

3.2 Reconnaissance automatique phonémique

Dans un premier temps, le nombre de phonèmes différents reconnus, sans tenir compte de leur nombre d'occurrences, a été étudié. En moyenne, 28,9 phonèmes différents sont reconnus sur les 32 phonèmes cibles (écart-type = 4,6, voir Figure 1). Un sujet n'a que trois phonèmes différents reconnus au cours de la production des 104 pseudo-mots attendus.

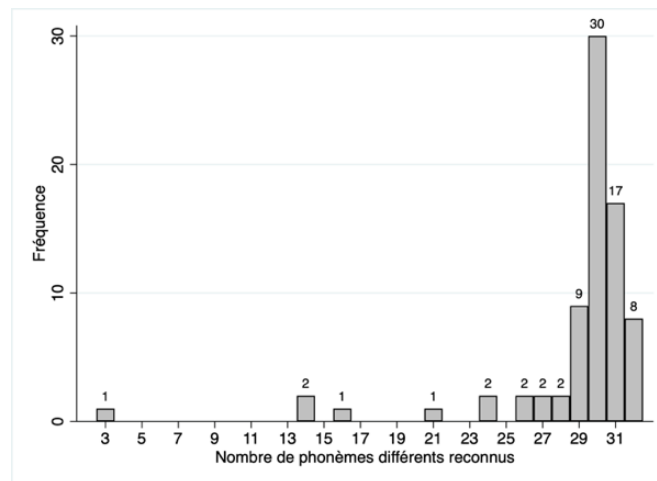


FIGURE 1 : Distribution du nombre de phonèmes différents reconnus pour chaque sujet par le système de reconnaissance automatique

Puis, nous avons dressé un inventaire phonémique global par sujet, en comptant toutes les occurrences de tous les phonèmes (cf. Table 1). Nous retrouvons alors une légère prépondérance des consonnes reconnues dans la parole des sujets par rapport aux voyelles, avec un taux moyen de consonnes à 50,9 % (la proportion de consonnes en français en population générale étant de 56,6 %, [Wioland, 1991](#)).

3.3 Liens articulatoires et moteurs

La table 2 représente les taux des différentes classes phonémiques significativement différents entre les sujets ne présentant pas d'anomalie (anomalie absente ou faible) et ceux présentant une anomalie lors de l'examen moteur (cf. section 2.3.1). Les classes « taux de consonnes bilabiales », « taux de consonnes alvéolaires », « taux de consonnes palatales » et « taux de consonnes orales » ne sont pas représentées, car elles ne montrent aucune différence significative, quelle que soit la localisation anatomique testée.

Consonnes	Moyenne (ET)	Minimum ; Maximum
csn	50,87 % (3,19)	33,33 % ; 60,98 %
occ	19,34 % (8,43)	0,00 % ; 40,00 %
bilabial	12,74 % (4,00)	0,00 % ; 24,00 %
labiodent	14,94 % (5,52)	0,00 % ; 32,47 %
alveol	38,50 % (9,44)	18,18 % ; 100,00 %
postalv	5,66 % (3,11)	0,00 % ; 13,53 %
palatal	21,75 % (5,67)	0,00 % ; 47,37 %
velaire	6,41 % (3,09)	0,00 % ; 13,01 %
sourde	24,60 % (8,07)	0,00 % ; 38,32 %
corale	86,68 % (5,87)	69,94 % ; 100,00 %

TABLE 1 : Tableau de distribution des taux de consonnes reconnues

		csn	occ	labiodent	postalv	velaire	sourde
Lèvres	Score A (p)	<i>N.S.</i>	<i>0,03</i>	<i>0,01</i>	<i>0,02</i>	<i>N.S.</i>	<i>0,049</i>
	0		20,17%	14,29%	6,03%		25,36%
	1		13,81%	19,26%	3,24%		19,53%
	Score F (p)	<i>N.S.</i>	<i>0,003</i>	<i>0,01</i>	<i>0,02</i>	<i>N.S.</i>	<i>N.S.</i>
	0		20,15%	14,30%	5,99%		
	1		13,25%	19,78%	3,24%		
Joues	Score A (p)	<i>N.S.</i>	<i>0,02</i>	<i>N.S.</i>	<i>N.S.</i>	<i>N.S.</i>	<i>0,001</i>
	0		20,31%				25,94%
	1		14,58%				17,99%
	Score F (p)	<i>0,02</i>	<i>0,006</i>	<i>N.S.</i>	<i>N.S.</i>	<i>N.S.</i>	<i>0,0002</i>
	0	51,56%	21,74%				27,44%
	1	49,85%	15,78%				20,40%
Mâchoire	Score A (p)	<i>N.S.</i>	<i>N.S.</i>	<i>N.S.</i>	<i>0,03</i>	<i>N.S.</i>	<i>N.S.</i>
	0				6,02%		
	1				3,73%		
	Score F (p)	<i>N.S.</i>	<i>N.S.</i>	<i>0,01</i>	<i>0,02</i>	<i>N.S.</i>	<i>N.S.</i>
	0			14,28%	6,03%		
	1			18,91%	3,45%		
Langue	Score A (p)	<i>0,004</i>	<i>0,0001</i>	<i>0,0002</i>	<i>N.S.</i>	<i>0,02</i>	<i>0,0001</i>
	0	51,30%	24,29%	12,20%		7,46%	28,98%
	1	50,60%	16,19%	16,68%		5,73%	21,80%
	Score F (p)	<i>0,001</i>	<i>0,0002</i>	<i>0,005</i>	<i>N.S.</i>	<i>0,04</i>	<i>0,001</i>
	0	51,41%	23,15%	12,94%		7,18%	27,94%
	1	50,46%	16,48%	16,44%		5,83%	22,10%
Vélopharynx	Score A (p)	<i>0,0475</i>	<i>0,01</i>	<i>N.S.</i>	<i>N.S.</i>	<i>N.S.</i>	<i>0,03</i>
	0	51,31%	21,06%				26,07%
	1	50,01%	15,97%				21,72%
	Score F (p)	<i>N.S.</i>	<i>N.S.</i>	<i>N.S.</i>	<i>N.S.</i>	<i>N.S.</i>	<i>0,01</i>
	0						26,46%
	1						20,95%

* 0 = pas d'anomalie, anomalie discrète ou modérée ; 1 = anomalie marquée, sévère ou quasi permanente ; N.S. : non significatif

* Score A = score de synthèse analytique ; Score F = score de synthèse fonctionnel. Le p correspond à la p-value du test de Mann-Whitney (entre anomalie absente ou faible et marquée).

Seules les proportions significativement différentes selon l'anomalie sont indiquées, avec le taux le plus faible indiqué en gras.

TABLE 2 : Moyenne des taux de consonnes reconnues par sujet, selon les scores moteurs analytiques et fonctionnels et par segment anatomique testé

Le taux de consonnes est ainsi significativement plus faible en cas d'atteinte de la langue, de même que le taux d'occlusives. Ce taux d'occlusives est également significativement plus faible en cas d'atteinte analytique ou fonctionnelle des joues et des lèvres. Le taux de consonnes sourdes est également moindre en cas d'atteinte analytique des lèvres, et d'atteinte analytique et fonctionnelle des joues, de la langue et du vélopharynx.

Dans le détail des catégories, aucune différence significative de proportion de taux de consonnes post-alvéolaires n'est retrouvée au niveau de la langue (en fonctionnel : anomalie absente ou faible à 5,56 % ; anomalie marquée à 5,74 %, $p=0,95$), alors que cette différence est retrouvée au niveau de la mâchoire sur la même classe phonémique (en fonctionnel : anomalie absente ou peu marquée à 6,03 % ; anomalie marquée à 3,45 %, $p=0,02$). Le taux de consonnes vélaires n'est significativement différent qu'en cas d'atteinte de la langue, avec un taux significativement plus important en cas d'absence d'anomalie. Enfin, en cas d'atteinte de la mâchoire, des lèvres et de la langue, le taux de consonnes labiodentales reconnues devient plus élevé.

4 Discussion

4.1 Résultats principaux

Cette étude montre l'effet que peut avoir une atteinte motrice, analytique ou fonctionnelle, sur la qualité de la production phonémique de la parole.

Les anomalies linguales décelées lors du bilan analytique sont ainsi liées à un taux plus faible de consonnes reconnues par un système de reconnaissance phonémique, d'occlusives, de postalvéolaires et de consonnes sourdes. Ces éléments sont cohérents avec l'altération des structures anatomiques en jeu dans la production de ce type de phonèmes. Seul le taux de consonnes post-alvéolaires reconnues n'est pas significativement associé à une altération de la langue, probablement lié au fait que les consonnes post-alvéolaires sont majoritairement des fricatives (/ʃ/, /ʒ/) et que le déficit est davantage lié à une atteinte de la mâchoire que l'on retrouve dans nos résultats. Enfin, une atteinte de la langue est également significativement associée à un taux plus important de consonnes labiodentales reconnues, pouvant être lié à un mécanisme de compensation où l'occlusion est délocalisée du niveau lingual (occlusion dégradée ou impossible) au niveau labio-dental.

Contrairement à [De Bruijn et al., 2012](#), nous ne retrouvons pas dans notre corpus de différence significative de taux de consonnes alvéolaires selon l'altération motrice, et ce quel que soit le segment anatomique. Ceci pourrait être dû à la composition de notre échantillon qui présente très majoritairement une atteinte orale (76 sujets, 30 sujets ayant une atteinte de l'oropharynx avec des sujets présentant une atteinte mixte orale et oropharyngée), ce qui ne permet pas de mettre en évidence ce type de dégradation.

De façon plus globale, nous retrouvons un taux d'occlusives significativement plus faible en cas d'anomalie analytique ou fonctionnelle au niveau des lèvres, des joues, de la langue et du vélopharynx (analytique seulement). En raison de la localisation de la pathologie tumorale et du traitement subi, majoritairement chirurgical (63/77, 82 %), le manque de mobilité des segments anatomiques combiné au défaut structurel limite la capacité d'occlusion des sujets qui ont alors tendance à produire des approximantes, reconnues comme fricatives par notre système de reconnaissance phonémique (taux d'occlusives reconnues plus faible). Le rôle du défaut est également retrouvé au niveau des anomalies analytiques et fonctionnelles labiales dans une population pour laquelle aucun sujet n'a présenté de localisation tumorale au niveau des lèvres. Ainsi,

les différences de taux d'occlusives, de labiodentales, postalvéolaires et sourdes retrouvées peuvent être liées à l'atteinte labiale secondaire relative aux conséquences du traitement chirurgical et radiothérapeutique, pouvant réduire la capacité de contention labiale ou justifier la mise en place d'une compensation plus locale labiodentale.

4.2 Limites de l'étude

Notre étude s'intéresse aux liens acoustico-moteurs après cancer oral ou oropharyngé, via la réalisation d'un inventaire phonémique automatique des consonnes. Toutefois, les recrutements étant encore en cours à l'heure de publication de ce papier, les résultats présentés ici ne peuvent être considérés que comme des tendances et non des résultats définitifs. Un recrutement d'un plus grand nombre de sujets permettra également de s'intéresser de façon plus précise aux liens pouvant exister entre la localisation de la pathologie, ses impacts analytiques et fonctionnels et ses répercussions sur la production de parole, notamment au niveau phonémique. Il est en effet possible que les résultats obtenus montrent des tendances différentes selon la région anatomique en jeu dans le cancer ou son traitement, notamment entre la zone orale et la zone oropharyngée.

De plus, notre étude s'est intéressée à l'impact du cancer oral ou oropharyngé sur les consonnes. L'observation du comportement d'un système de reconnaissance phonémique au niveau des voyelles pourrait ainsi permettre de compléter l'analyse, notamment via les effets de modification des volumes intra-oraux induits principalement par le traitement chirurgical.

4.3 Perspectives

L'étude des liens acoustico-moteurs est importante en cancérologie ORL. Un lien entre la production de parole et le score moteur permettrait de simplifier et d'optimiser les évaluations cliniques chez des patients souvent fatigables, en ne proposant aux sujets que de réaliser une tâche de répétition de pseudo-mots. C'est par une analyse automatique des enregistrements de ces productions que seraient déterminées à la fois les unités linguistiques déficitaires, mais également les altérations dynamiques habituellement retrouvées lors d'un bilan analytique orthophonique complet.

Ce type d'analyse permettrait également de fournir aux cliniciens des éléments de compréhension plus fins des mécanismes en jeu dans les liens articulatoires et acoustiques de bas niveau de la parole. Les stratégies thérapeutiques de patients seront améliorées, grâce à une combinaison d'approches ciblant l'intelligibilité de la parole et d'autres plus écologiques ciblant la communication et la compréhensibilité.

5 Conclusion

L'utilisation de systèmes de reconnaissance automatique de phonèmes aboutissant à la réalisation d'un inventaire phonémique par sujet permet de mettre en évidence une influence de la pathologie cancérologique orale ou oropharyngée et de ses traitements sur la reconnaissance de phonèmes. Notamment, l'altération motrice de la langue a un impact large sur la reconnaissance des consonnes (occlusives et vélaires notamment), et certains mécanismes de compensation ont pu être retrouvés (avec des taux de reconnaissance de consonnes labiodentales plus élevés en cas d'atteinte des lèvres, de la mâchoire ou de la langue). Cette étude doit donc se poursuivre par une analyse de la reconnaissance automatique des voyelles, et la mise en perspective de ces résultats avec les données cliniques et de traitement des sujets.

Remerciements

Cette étude a bénéficié d'un financement du ministère de la Santé (PHRIP, 2019, PHRIP-19-0004).

Références

- AUZOU, P., & ROLLAND-MONNOURY, V. (2019). BECD : Batterie d'Évaluation Clinique de la Dysarthrie (*Ortho Édit*).
- CHRISTENSEN, H., CUNNINGHAM, S., FOX, C., GREEN, P., & HAIN, T. (2012). A comparative study of adaptive, automatic recognition of disordered speech. *Interspeech 2012*, 2, 1776–1779. <https://doi.org/10.21437/Interspeech.2012-484>
- DE BRUIJN, M. J., BOSCH, L. TEN, KUIK, D. J., WITTE, B. I., LANGENDIJK, J. A., RENE LEEMANS, C., & VERDONCK-DE LEEUW, I. M. (2012). Acoustic-phonetic and artificial neural network feature analysis to assess speech quality of stop consonants produced by patients treated for oral or oropharyngeal cancer. *Speech Communication*, 54(5), 632–640. <https://doi.org/10.1016/j.specom.2011.06.005>
- DOYLE, P. C., LEEPER, H. A., KOTLER, A. L., THOMAS-STONELL, N., O'NEILL, C., DYLKE, M. C., & ROLLS, K. (1997). Dysarthric speech: a comparison of computerized speech recognition and listener intelligibility. *Journal of Rehabilitation Research and Development*, 34(3), 309–316. <http://www.ncbi.nlm.nih.gov/pubmed/9239624>
- FEX, S. (1992). Perceptual evaluation. *Journal of Voice*, 6(2), 155–158.
- GHIO, A., GIUSTI, L., BLANC, E., PINTO, S., LALAIN, M., ROBERT, D., FREDOUILLE, C., & WOISARD, V. (2016). Quels tests d'intelligibilité pour évaluer les troubles de production de la parole ? *Journées d'Étude Sur La Parole*, 589–596. <https://hal.archives-ouvertes.fr/hal-01372037>
- GHIO, A., LALAIN, M., GIUSTI, L., POUCHOUIN, G., ROBERT, D., REBOURG, M., FREDOUILLE, C., LAARIDH, I., & WOISARD, V. (2018). Une mesure d'intelligibilité par décodage acoustico-phonétique de pseudo-mots dans le cas de parole atypique. *XXXIIe Journées d'Études Sur La Parole*, 285–293. <https://doi.org/10.21437/JEP.2018-33>
- GHIO, A., LALAIN, M., REBOURG, M., MARCZYK, A., FREDOUILLE, C., & WOISARD, V. (2022). Validation of an Intelligibility Test Based on Acoustic-Phonetic Decoding of Pseudo-Words: Overall Results from Patients with Cancer of the Oral Cavity and the Oropharynx. *Folia Phoniatica et Logopaedica*, 74(3), 209–222. <https://doi.org/10.1159/000519427>
- LAPOTRE-LEDOUX, B., REMONTET, L., UHRY, Z., DANTONY, E., GROSCLAUDE, P., MOLINIE, F., WORONOFF, A.-S., LECOFFRE-BERNARD, C., LAFAY, L., DEFOSSEZ, G., D'ALMEIDA, T., & FRANCIM, R. français des registres de cancers. (2023). Incidence Des Principaux Cancers En France Métropolitaine En 2023 Et Tendances Depuis 1990 / Main Cancers Incidence in Metropolitan France in 2023 and Trends Since 1990. *Bulletin Épidémiologique Hédomadaire*, 12–13, 188–204. http://beh.santepubliquefrance.fr/beh/2023/12-13/2023_12-13_1.html
- LAZARUS, C. L., HUSAINI, H., ANAND, S. M., JACOBSON, A. S., MOJICA, J. K., BUCHBINDER, D., & URKEN, M. L. (2013). Tongue Strength as a Predictor of Functional Outcomes and Quality of Life after Tongue Cancer Surgery. *Annals of Otology, Rhinology & Laryngology*, 122(6), 386–397. <https://doi.org/10.1177/000348941312200608>
- MAIER, A., HADERLEIN, T., STELZLE, F., NÖTH, E., NKENKE, E., ROSANOWSKI, F., SCHÜTZENBERGER, A., & SCHUSTER, M. (2010). Automatic Speech Recognition Systems for the Evaluation of Voice and Speech Disorders in Head and Neck Cancer. *EURASIP Journal on Audio, Speech, and Music Processing*, 2010(April 2014), 1–7. <https://doi.org/10.1155/2010/926951>
- MIDDAG, C. (2013). Automatische analyse van pathologische spraak Automatic Analysis of Pathological Speech.

- MIDDAG, C., VAN NUFFELEN, G., MARTENS, J. P., & DE BODT, M. (2008). Objective intelligibility assessment of pathological speakers. *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH*, 1745–1748.
- MLYNAREK, A., RIEGER, J., HARRIS, J., O'CONNELL, D., AL-QAHTANI, K., ANSARI, K., CHAU, J., & SEIKALY, H. (2008). Methods of functional outcomes assessment following treatment of oral and oropharyngeal cancer: review of the literature. *Journal of Otolaryngology - Head & Neck Surgery*, 37(1), 2–10. <https://doi.org/10.2310/7070.2008.1001>
- PLISSON, L., PILLOT-LOISEAU, C., & CREVIER-BUCHMAN, L. (2017). Intelligibilité de la parole après le traitement d'un cancer de l'oropharynx : étude descriptive chez sept patients en pré-traitement et en post-traitement précoce. *7èmes Journées de Phonétique Clinique (JPC7)*.
- POMMEE, T., BALAGUER, M., MAUCLAIR, J., PINQUIER, J., & WOISARD, V. (2021). Intelligibility and comprehensibility: A Delphi consensus study. *International Journal of Language & Communication Disorders*, 1–44. <https://doi.org/10.1111/1460-6984.12672>
- POMMEE, T., BALAGUER, M., MAUCLAIR, J., PINQUIER, J., & WOISARD, V. (2022). Assessment of adult speech disorders: current situation and needs in French-speaking clinical practice. *Logopedics Phoniatrics Vocology*, 47(2), 92–108. <https://doi.org/10.1080/14015439.2020.1870245>
- REICH, M. (2009). Cancer et image du corps : identité, représentation et symbolique : Le corps retrouvé (French). *Cancer and Body Image : Identity, Representation* (English), 85(3), 247–254. <https://doi.org/10.3917/inpsy.8503.0247>
- WIOLAND, F. (1991). *Prononcer les mots du français – Des sons et des rythmes*. Hachette. ISBN : 2-01-017482-8