



HAL
open science

Paramètres perceptifs expliquant la sévérité du trouble de parole mesurée automatiquement en cancérologie ORL

Mathieu Balaguer, Timothy Pommée, Jérôme Farinas, Julien Pinquier,
Virginie Woisard

► To cite this version:

Mathieu Balaguer, Timothy Pommée, Jérôme Farinas, Julien Pinquier, Virginie Woisard. Paramètres perceptifs expliquant la sévérité du trouble de parole mesurée automatiquement en cancérologie ORL. Rééducation orthophonique, 2021, Chapitre: "De l'exploration à la prise en soins de la voix chez l'adulte: données actuelles.. sur la voie des voix", 286, pp.1-13. hal-03277422

HAL Id: hal-03277422

<https://ut3-toulouseinp.hal.science/hal-03277422v1>

Submitted on 5 Jul 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.

Paramètres perceptifs expliquant la sévérité du trouble de parole mesurée automatiquement en cancérologie ORL

MATHIEU BALAGUER

Orthophoniste, Doctorant en Informatique et télécommunications
IRIT, CNRS, Université Paul Sabatier Toulouse III, 118 Route de Narbonne
31062 Toulouse Cedex 9
CHU Larrey, Toulouse
mathieu.balaguer@irit.fr

TIMOTHY POMMEE

Orthophoniste, Doctorant en Informatique et télécommunications
IRIT, CNRS, Université Paul Sabatier Toulouse III, 118 Route de Narbonne
31062 Toulouse Cedex 9
timothy.pommee@irit.fr

JEROME FARINAS

Maître de Conférence des Universités
IRIT, CNRS, Université Paul Sabatier Toulouse III, 118 Route de Narbonne
31062 Toulouse Cedex 9
jerome.farinas@irit.fr

JULIEN PINQUIER

Maître de Conférence des Universités
IRIT, CNRS, Université Paul Sabatier Toulouse III, 118 Route de Narbonne
31062 Toulouse Cedex 9
julien.pinquier@irit.fr

VIRGINIE WOISARD

Médecin ORL, Phoniatre
CHU Larrey, 24 chemin de Pouvoirville
31059 Toulouse Cedex 9
woisard.v@chu-toulouse.fr

RESUME

Introduction : Les cancers de la cavité buccale et de l'oropharynx atteignent les structures anatomiques en jeu dans la production de la parole. L'évaluation orthophonique de la parole est actuellement principalement basée sur des évaluations perceptives, présentant de nombreuses limites en termes de fiabilité. Le projet C2SI (Carcinologic Speech Severity Index) a permis de construire un score automatique, fiable et objectif de sévérité du trouble de parole chez ces patients.

Objectif : Déterminer quels paramètres de parole impactent la sévérité du trouble de parole.

Méthode : 87 sujets traités pour un cancer de la cavité buccale ou de l'oropharynx ont été inclus. Le score automatique de sévérité C2SI a été calculé pour ces sujets. Un jury de six auditeurs experts a évalué perceptivement la voix, la résonance, la prosodie et la production phonémique.

Résultats : L'évaluation perceptuelle des paramètres de parole montre un accord inter-juges moyen. Une analyse en régression linéaire réalisée sur l'ensemble des sujets montre l'impact de la prosodie ($p=0,02$), de la résonance ($p<0,001$) et de la production phonémique ($p=0,001$) sur la sévérité du trouble ($R^2=0,75$). Toutefois, une analyse stratifiée par localisation montre que si la résonance est un paramètre expliquant la sévérité du trouble quelle que soit la localisation ($p<0,01$), la production phonémique contribue à l'explication en cas de tumeur orale ($p<0,01$), et la qualité vocale ($p=0,03$) en cas de tumeur oropharyngée.

Discussion – Conclusion : L'influence des différents paramètres de parole sur la sévérité du trouble diffère selon la localisation tumorale. Une automatisation de l'extraction de ces paramètres permettrait une meilleure fiabilité de l'évaluation, pour adapter les stratégies thérapeutiques au plus près des besoins des patients, la mesure de l'impact du trouble sur leur communication restant à explorer.

MOTS-CLES : parole, évaluation perceptuelle, évaluation automatique, cancer

Perceptual parameters explaining the severity of speech impairment measured automatically in ENT oncology

ABSTRACT

Introduction: Oral or oropharyngeal cancer affect the anatomical structures involved in speech production. Currently, speech assessment is mainly based on perceptual measurements, with many limitations in terms of reliability. The C2SI (Carcinologic Speech Severity Index) project allowed to build an automatic, reliable and objective score for the severity of the speech disorder in these patients.

Objective: To determine which specific speech parameters impact the severity of the speech disorder.

Method: 87 subjects treated for oral or oropharyngeal cancer were included. The automatic C2SI severity score was calculated for these subjects. A panel of six expert listeners perceptually assessed voice, resonance, prosody and phonemic production.

Results: The perceptual evaluation of the specific speech parameters showed a moderate inter-judge agreement. A linear regression analysis performed on all subjects shows the impact of prosody ($p=.02$), resonance ($p<.001$) and phonemic production ($p=.001$) on the severity of the disorder ($R^2=.75$). However, a stratified analysis by tumor location shows that if resonance explains the severity of the disorder whatever the location ($p<.01$), phonemic production contributes to the explanation only in case of oral tumor ($p<.01$), and prosody ($p=.04$) and voice ($p=.03$) only in case of oropharyngeal tumor.

Discussion - Conclusion: The influence of the different specific speech parameters on the severity of the disorder differs according to the tumor location. An automatic extraction of these parameters would allow a better reliability of the evaluation, to adapt the therapeutic strategies closer to the needs of the patients. However, the measurement of the impact of the disorder on their communication abilities remains to be explored.

KEY WORDS: speech, perceptual assessment, automatic assessment, cancer

◆ INTRODUCTION

Les cancers de la cavité buccale et de l'oropharynx ont une incidence élevée en France¹ : plus de 13 500 nouveaux cas ont été déclarés en 2018, ce qui en fait le troisième cancer le plus fréquent. Environ trois quarts des sujets sont des hommes, avec 18,3 cas pour 100 000 personnes-années (taux standardisé mondial, contre 5,8 chez les femmes).

En raison de leur localisation (1,2), les cancers de la cavité buccale et de l'oropharynx affectent les capacités de production de parole. De plus, les traitements consécutifs à ces cancers auront également des conséquences fonctionnelles. Différentes études ont montré que la taille de la tumeur (3–5) et le type de traitement (chirurgie et/ou radiothérapie et/ou chimiothérapie) (6) déterminent la qualité de la production de parole. Or, la dégradation de la parole aura un impact majeur sur les fonctions de communication des patients, ainsi que sur leur qualité de vie (3,7,8) et leur dynamique sociale (9). En considérant également la diminution de la mortalité (de 13,4 pour 100 000 en 1990 à 4,9 en 2018), et l'augmentation de l'espérance de vie consécutives au traitement des cancers oraux ou oropharyngés, il est désormais indispensable de s'intéresser à la prise en charge des troubles de parole, en raison de leur impact sur les limitations fonctionnelles d'activité et les restrictions psychosociales des patients.

La stratégie thérapeutique pour les patients atteints d'un cancer oral ou oropharyngé doit donc être définie sur la base d'une évaluation spécifique du trouble de la parole. À l'heure actuelle, l'évaluation clinique, qui est principalement réalisée par des orthophonistes, reste essentiellement perceptive (10,11). Les thérapeutes s'appuient sur différentes méthodes de mesure pour quantifier la perception de la parole : échelles visuelles analogiques, échelles de Likert, estimations directes de la magnitude (12,13), transcription du stimulus (14), ou encore évaluation de la compréhension de la parole (15). Parmi toutes ces valeurs, le score de sévérité du trouble de la parole perceptif sur une tâche de description d'image est le score le plus informatif (16). Toutefois, l'évaluation perceptive présente de nombreuses limites en termes de fiabilité. En effet, l'évaluation quantitative d'un trouble de parole est variable selon le juge qui l'évalue (effets d'expertise, référents internes différents...) (17), mais un même juge peut attribuer des scores différents selon le contexte d'évaluation, sa disponibilité mentale, son habitude à la parole pathologique (18)... L'obtention d'un score issu d'un traitement automatique du signal de parole, champ de recherche en plein développement (19–21), permet de pallier ces biais.

Le trouble de la parole peut être défini à différents niveaux d'analyse (22,23). Les déficits de bas niveau (au niveau du phonème ou de la syllabe) affectent l'intelligibilité de la parole et réduisent la capacité à interagir avec les autres (24). L'intelligibilité est autrement définie comme la proportion de la parole comprise (25) ou le taux de mots correctement transcrits (15). À un niveau plus élevé (comme au niveau des mots ou des phrases), les indices contextuels syntaxiques et sémantiques permettent à l'auditeur de reconstruire l'information véhiculée par le discours (26). Ces phénomènes de compensation sont essentiels pour transmettre un message informatif dans le contexte d'une parole dégradée ou pathologique (27). Ainsi, la sévérité du trouble peut être définie comme une notion plus globale, où divers éléments du signal vocal sont pris en compte : la qualité du code acoustico-phonétique, le débit de parole et d'autres paramètres temporels et/ou prosodiques relatifs à la gêne perçue (12,28). La nature globale de la sévérité du trouble de la parole, supposée dans sa définition, devrait cependant être explorée plus précisément afin de comprendre les paramètres perceptifs spécifiques qui sous-tendent son évaluation.

¹ Rapport Santé publique France, juillet 2019 : <https://www.santepubliquefrance.fr/maladies-et-traumatismes/cancers/cancer-du-sein/documents/rapport-synthese/estimations-nationales-de-l-incidence-et-de-la-mortalite-par-cancer-en-france-metropolitaine-entre-1990-et-2018-volume-1-tumeurs-solides-etud>

L'objectif de cette étude est de déterminer quels paramètres perceptifs spécifiques de parole impactent la sévérité du trouble de parole chez les patients traités pour un cancer oral ou oropharyngé.

◆ **MATERIEL ET METHODES**

Base de données C2SI

Cette étude est basée sur le corpus du projet C2SI (Carcinologic Speech Severity Index) (21), financé par l'Institut National du Cancer (Subvention INCa SHS n°2015-135). L'objectif de ce projet était d'obtenir une mesure de l'impact des traitements du cancer des voies aérodigestives supérieures (cavité buccale et pharynx) sur la production vocale par traitement automatique de la parole, en comparaison avec les méthodes perceptives.

Cette base de données comprend les informations issues des consultations de suivi post-traitement des patients atteints de cancer des VADS, datées entre 2015 et 2016 à l'Institut Universitaire du Cancer de Toulouse - Oncopole. Pour être inclus, les patients devaient avoir terminé leur protocole thérapeutique depuis au moins six mois, être en rémission clinique et ne pas présenter de trouble de la parole d'autre origine (ex : bégaiement).

Échantillons de parole

Les patients recrutés dans le cadre du projet C2SI ont vu leur parole enregistrée lors d'une tâche de description d'image. Cet enregistrement a été réalisé le jour de leur consultation de suivi, dans une salle anéchoïque.

Les sujets devaient décrire une image choisie au hasard parmi un ensemble de représentations de paysages de bord de mer avec un bateau. L'échantillon de parole devait durer au moins une minute, afin de disposer de suffisamment de données pour être exploitable.

La production de la parole a été réalisée à l'aide d'un enregistreur numérique avec une fréquence d'échantillonnage de 48 kHz et une résolution de 16 bits, en mode mono. Les fichiers sonores ont été enregistrés au format PCM linéaire (exporté en WAVE).

Jury d'écoute

Six auditeurs experts (cinq orthophonistes et un phoniatre travaillant dans un service hospitalier d'oncologie ORL) ont évalué les échantillons de parole à l'aide d'un casque dans un environnement calme.

À l'aide d'un formulaire dédié, ils ont évalué quatre paramètres perceptifs spécifiques de la parole (qualité de la voix, résonance, prosodie et altérations phonémiques) sur une échelle allant de zéro (aucune déficience) à trois (déficience grave). Ces quatre paramètres n'étaient pas explicitement définis dans les instructions données aux auditeurs. Suite à la cotation de ces quatre paramètres, ces auditeurs ont également évalué perceptivement la sévérité du trouble de parole de façon plus globale, sur une échelle de zéro (trouble majeur) à dix (absence de trouble).

Score automatique C2SI

L'étude C2SI a permis d'aboutir à la construction d'un score automatique de sévérité (29), consécutif à la modélisation du score perceptif de sévérité attribué par un jury d'écoute au moyen de différents paramètres extraits par un traitement automatique du signal de parole. Ces paramètres correspondent à des indices acoustiques relatifs à différents niveaux d'analyse du signal de parole (variabilité de la fréquence fondamentale, variation de la hauteur, scores automatiques de vraisemblance phonémique), pour aboutir à une corrélation forte à 0,87 entre le score automatique et le score perceptif de sévérité.

Dans le cadre de cette étude, nous choisirons de nous intéresser à l'analyse de ce score automatique de sévérité. En effet, en raison de sa construction au moyen du score perceptif de sévérité, il porte une information globale concernant le trouble de la parole, mais il est aussi plus fiable que les scores perceptifs classiques car il est débarrassé des phénomènes de variabilité individuelle. Ce score C2SI est construit sur une échelle de zéro (correspondant à une sévérité majeure du trouble) à dix (trouble considéré comme non sévère).

Analyses statistiques

Une analyse statistique descriptive a été effectuée dans un premier temps. Concernant les paramètres perceptifs spécifiques, l'accord inter-juges a été mesuré en calculant un coefficient de corrélation intra-classe (ICC) à deux facteurs à effet aléatoire.

L'analyse de la sévérité du trouble de la parole en fonction des paramètres perceptifs spécifiques a été réalisée en deux étapes : une analyse bivariée de chaque paramètre avec la sévérité (corrélations de Spearman non paramétriques), et une analyse multivariée utilisant des analyses de régression linéaire robustes.

Toutes les analyses ont été effectuées à l'aide du logiciel Stata 14.1.

◆ RESULTATS

Description de la base de données C2SI

La base de données C2SI comprend 87 sujets. Tous ont été inclus dans cette étude : 51 hommes et 36 femmes, avec un âge moyen de 65,8 ans (écart-type : 9,6).

Trente-cinq patients avaient été traités pour un cancer de la cavité buccale et 52 pour un cancer de l'oropharynx. Le traitement était uniquement chirurgical pour cinq patients, en radiothérapie seule pour 14 patients, et la chirurgie associée à la radiothérapie pour 68 patients.

Enfin, le délai médian depuis la fin du traitement était de 39 mois (intervalle interquartile : [21, 91]).

Analyse descriptive des scores

Le score automatique C2SI a pu être calculé au final sur les 59 sujets ne disposant pas de donnée manquante (32 hommes et 27 femmes, d'âge moyen 65,2 ans – SD=10,17). Parmi eux, 27 avaient été traités pour un cancer de la cavité buccale et 32 pour un cancer de l'oropharynx. 51 ont été traités chirurgicalement parmi lesquels 47 ont également été traités en radiothérapie. Les huit patients restants ont été traités uniquement par radiothérapie. Ainsi, sur les 59 patients, le score C2SI est en moyenne de 6,10/10 (SD=1,83 – pour rappel, le score C2SI évolue dans le sens opposé des paramètres perceptifs spécifiques : plus le score C2SI est faible, plus la sévérité du trouble est importante).

Concernant les paramètres perceptifs spécifiques, les résultats sont présentés dans le Tableau 1. La moyenne générale de l'évaluation des altérations phonémiques est la plus élevée ($m=1,47$, $SD=0,87$), tandis que la prosodie est le paramètre ayant le score le plus bas ($m=0,62$, $SD=0,52$). À l'exception de l'auditeur 3, les altérations phonémiques présentent le score le plus élevé parmi les paramètres perceptifs pour tous les évaluateurs, donc le paramètre le plus altéré. L'analyse par auditeur montre que si deux d'entre eux (auditeurs 1 et 6) sous-estiment l'atteinte de ces paramètres par rapport aux autres (avec des scores systématiquement inférieurs à la moyenne générale), un (auditeur 3) les surestime (en moyenne $+0,49$ point).

Tableau 1 Moyennes et écarts-types (entre parenthèses) des paramètres perceptifs, par auditeur

	Qualité vocale	Résonance	Prosodie	Altérations phonémiques
Global	1,06 (0,64)	1,12 (0,71)	0,62 (0,52)	1,47 (0,87)
1	0,85 (0,90)	0,88 (1,01)	0,59 (0,83)	1,09 (1,08)
	-0,21	-0,26	-0,03	-0,39
2	0,59 (0,84)	1,08 (0,89)	0,68 (0,80)	1,61 (0,94)
	-0,47	-0,07	0,07	0,12
3	1,76 (0,82)	1,68 (0,90)	1,22 (0,85)	1,59 (1,04)
	0,70	0,54	0,60	0,11
4	1,36 (0,86)	1,39 (0,94)	0,22 (0,54)	1,54 (1,01)
	0,30	0,24	-0,39	0,06
5	1,04 (0,81)	0,79 (0,93)	0,44 (0,65)	1,62 (1,00)
	-0,02	-0,35	-0,18	0,13
6	0,75 (0,78)	1,04 (0,95)	0,55 (0,74)	1,44 (0,87)
	-0,30	-0,11	-0,07	-0,04

Les valeurs grisées correspondent aux différences entre la moyenne du paramètre évalué par l'auditeur et la moyenne générale du paramètre parmi les auditeurs (gris foncé : surestimation du trouble ; gris clair : sous-estimation).

Accord inter-juges sur les paramètres perceptifs spécifiques

Afin de déterminer l'accord entre les différents auditeurs, un ICC a été calculé (Tableau 2). Les ICC pour la prosodie et la qualité de la voix sont faibles, la valeur pour la résonance est acceptable, et seul l'ICC pour les distorsions phonémiques est considéré comme bon.

Tableau 2 ICC pour chaque paramètre perceptif spécifique, avec intervalle de confiance à 95 %

	ICC	Bornes de l'intervalle de confiance à 95 %	Interprétation (30)
Qualité vocale	0,37	[0,27 ; 0,47]	Faible
Résonance	0,42	[0,32 ; 0,52]	Correct
Prosodie	0,30	[0,20 ; 0,39]	Faible

Altérations phonémiques	0,67	[0,59 ; 0,75]	Bon
-------------------------	------	---------------	-----

Interprétation (30): en gris clair, ICC faible ; en gris intermédiaire, ICC moyen ; en gris foncé, ICC bon

En raison des faibles ICC pour les paramètres perceptifs spécifiques (Tableau 2) et de l'écart entre les moyennes des auditeurs individuels et les moyennes globales (Tableau 1), des limites d'agrément par paramètre ont été fixées pour étudier si chaque auditeur dépasse ces limites. Les limites d'agrément sont définies comme suit : [(moyenne des autres auditeurs) - 1.96SD, (moyenne des autres auditeurs) + 1.96SD]. Par exemple, pour examiner si l'auditeur 1 est dans les limites d'agrément, les limites suivantes ont été calculées : [(moyenne des auditeurs 2 à 6) - 1.96SD, (moyenne des auditeurs 2 à 6) + 1.96SD]. Les résultats de ces calculs sont présentés dans le Tableau 3.

Tableau 3 Comparaison du score moyen attribué par chaque auditeur avec la moyenne des scores des autres évaluateurs, pour chaque paramètre

Juge	Qualité vocale			Résonance			Prosodie			Altérations phonémiques		
	Valeur	m - 1,96 SD	m + 1,96 SD	Valeur	m - 1,96 SD	m + 1,96 SD	Valeur	m - 1,96 SD	m + 1,96 SD	Valeur	m - 1,96 SD	m + 1,96 SD
1	0,85	0,28	1,92	0,88	0,60	1,80	0,59	-0,03	1,28	1,09	1,43	1,68
2	0,59	0,43	1,87	1,08	0,51	1,81	0,68	-0,05	1,26	1,61	1,08	1,83
3	1,76	0,40	1,44	1,68	0,64	1,44	1,22	0,19	0,81	1,59	1,08	1,84
4	1,36	0,20	1,79	1,39	0,49	1,70	0,22	0,16	1,23	1,54	1,08	1,86
5	1,04	0,21	1,91	0,79	0,66	1,77	0,44	0,02	1,28	1,62	1,08	1,83
6	0,75	0,32	1,91	1,04	0,52	1,81	0,55	-0,02	1,28	1,44	1,09	1,88

En gris, les scores qui se situent en dehors des limites d'agrément.

Les scores moyens de l'auditeur 3 sont hors des limites d'agrément pour trois des quatre paramètres (qualité de la voix, résonance et prosodie), avec une surestimation systématique de ces scores. Les scores moyens de l'auditeur 1 sont hors limites d'agrément pour les altérations phonémiques, avec une sous-estimation de ce score.

Les ICC ont donc été recalculés dans différentes configurations, afin d'analyser si l'élimination d'un ou des deux auditeurs marginaux (1 et 3) améliore l'accord inter-juges. Même lorsque les auditeurs périphériques sont retirés du panel, l'amélioration des ICC reste faible (Tableau 4).

Tableau 4 ICC pour chacun des paramètres perceptifs en utilisant différents panels d'auditeurs.

	Global		Juge 3 exclus		Juges 1 et 3 exclus	
	ICC	Intervalle de confiance à 95 %	ICC	Intervalle de confiance à 95 %	ICC	Intervalle de confiance à 95 %
Qualité vocale	0,37	[0,27 ; 0,47]	0,45	[0,34 ; 0,55]	0,46	[0,35 ; 0,58]
Résonance	0,42	[0,32 ; 0,52]	0,47	[0,37 ; 0,57]	0,44	[0,32 ; 0,55]

Prosodie	0,30	[0,20 ; 0,39]	0,39	[0,28 ; 0,49]	0,31	[0,19 ; 0,43]
Altérations phonémiques	0,67	[0,59 ; 0,75]	0,67	[0,59 ; 0,76]	0,76	[0,68 ; 0,83]

Blanc : ICC faible ; gris clair : ICC moyen ; gris intermédiaire : ICC bon ; gris foncé : ICC excellent (30)

Explication du score de sévérité des troubles de la parole par les paramètres perceptifs

Une analyse bivariée entre le score C2SI de sévérité du trouble de la parole et chaque score des quatre paramètres perceptifs a été réalisée, au moyen des coefficients de corrélation de Spearman. La qualité de la voix ($r = -0,46$) et la prosodie ($r = -0,55$) sont modérément corrélées avec le score global de gravité du trouble, tandis que la résonance ($r = -0,75$) et les distorsions phonémiques ($r = -0,73$) sont fortement corrélées avec le score C2SI (31).

Des analyses en régression linéaire robuste ont été utilisées pour modéliser la sévérité du trouble de la parole C2SI en fonction des paramètres perceptifs spécifiques. Une première analyse a été réalisée sur l'ensemble des patients, puis deux analyses stratifiées ont été effectuées en fonction de la localisation de la tumeur (Tableau 5).

Tableau 5 Résultat de l'analyse en régression linéaire robuste

	Constante	Qualité vocale	Résonance	Prosodie	Altérations phonémiques	R ²
Tous	9,43	-0,35	-1,10 *	-0,92 *	-0,76 *	0,75
CB seuls	9,02	0,16	-1,39 *	-0,92	-0,81 *	0,77
OP seuls	9,78	-0,95 *	-1,21 *	-0,82	-0,47	0,79

CB : patients traités pour un cancer de la cavité buccale ; OP : patients traités pour un cancer de l'oropharynx ; une étoile indique une valeur statistiquement significative NS : non significatif

L'analyse en régression réalisée sur l'ensemble des sujets montre que si la qualité vocale ne semble pas en jeu dans l'explication de la sévérité du trouble de la parole, tous les autres paramètres perceptifs (résonance, prosodie et altérations phonémiques) sont des prédicteurs significatifs de la sévérité globale. Par contre, en analysant selon la localisation, seuls deux paramètres influencent la sévérité du trouble de parole. Un des deux est systématiquement la résonance, quelle que soit la localisation. L'autre est les altérations phonémiques pour la cavité buccale, et la qualité vocale pour l'oropharynx.

◆ DISCUSSION

Paramètres perceptifs impliqués dans l'évaluation de la gravité des troubles de la parole

L'analyse multivariée a montré que la qualité de la voix, la résonance, la prosodie et les altérations phonémiques sont toutes prises en compte lors de la détermination d'une évaluation globale de la sévérité des troubles de la parole chez les patients atteints de cancer de la cavité buccale ou de l'oropharynx. Toutefois, l'analyse stratifiée par localisation tumorale montre des résultats différents.

La qualité vocale n'est pas significativement prise en compte dans l'évaluation de la sévérité globale du trouble de parole chez les patients traités pour un cancer de la cavité buccale. Cela peut s'expliquer par l'anatomie et la physiologie des voies aérodigestives supérieures, où la production phonatoire se

déroule dans la zone laryngo-pharyngée et non dans la cavité buccale. Ces résultats sont cohérents avec l'anatomie et la physiologie des voies aérodigestives supérieures, la résonance étant souvent définie comme étant liée aux troubles vélopharyngés (32). De la même manière, l'absence de significativité de la prosodie chez ces patients peut être mise en lien avec le fait que les paramètres prosodiques en français sont en grande partie relatifs à la fréquence fondamentale ou à l'intensité (33) dont la source est laryngée.

L'absence de significativité des altérations phonémiques chez les patients traités pour un cancer de l'oropharynx est congruente avec les données de la littérature (2). Même si ces altérations concernent le principal déficit en termes de production de parole dans notre population d'étude, les altérations phonémiques atteignent majoritairement les patients traités pour une tumeur de la cavité buccale (34,35).

La significativité des quatre paramètres chez tous les auditeurs peut s'expliquer par deux éléments : les instructions de notation et la définition de la sévérité du trouble. Les auditeurs devaient évaluer les quatre paramètres spécifiques avant de noter la sévérité perçue. Ce score perceptif de sévérité n'a pas été utilisé dans notre étude de façon directe, mais le score automatique C2SI a été modélisé sur le score de sévérité perceptive. Ainsi, le processus de notation peut avoir induit un biais d'évaluation, l'évaluation des paramètres perceptifs pouvant influencer l'évaluation ultérieure de la sévérité, qui a ensuite pu se répercuter sur le score automatique suite à la modélisation. D'autre part, la sévérité du trouble était définie comme le « degré d'altération globale du signal », ce qui pourrait inciter à prendre en compte de façon globale l'ensemble des informations contenues dans le signal vocal et de parole pour évaluer la sévérité du trouble.

À l'exception des scores d'altérations phonémiques, l'accord inter-juges n'était que faible à modéré. De plus, l'analyse des limites d'agrément montre que certains juges ont tendance à systématiquement évaluer de façon plus sévère certains paramètres. Ces deux éléments d'analyse tendent à montrer qu'il existe des différences dans la façon d'appréhender la mesure perceptive d'un trouble, même au sein d'un panel de six juges experts. Ces dissemblances pourraient s'expliquer par des stratégies de notation sous-jacentes idiosyncrasiques entre les auditeurs. Dans ce contexte, il semblerait pertinent d'étudier les facteurs en jeu dans le profil d'évaluateur de chaque juge. L'analyse des éléments de contexte d'évaluation, de disponibilité mentale ou de familiarisation à la parole pathologique (18) au moyen d'échelles visuelles analogiques auto-rapportées par les juges, ainsi que d'autres paramètres tels que l'ancienneté professionnelle auprès des patients en cancérologie par exemple pourrait permettre d'aboutir à une sorte de pondération du score attribué par les juges. Cette démarche tendrait à homogénéiser les évaluations perceptives entre juges notamment dans le cadre du suivi des patients, avec des scores comparables d'un moment à l'autre du suivi.

Variabilité des scores perceptifs et complémentarité perceptif / automatique

À l'exception des scores d'altérations phonémiques, les paramètres spécifiques n'ont montré qu'un ICC faible à modéré. Cet accord limité entre les auditeurs exprime des différences dans la notation des paramètres perceptifs, même entre experts. Cet accord limité entre juges peut s'expliquer par le fait que les concepts sous-tendus par les paramètres spécifiques n'ont pas été définis au préalable. Ainsi, chaque auditeur a dû s'appuyer sur leurs propres référents internes pour estimer chaque paramètre. Si les altérations phonémiques sont facilement identifiables et communément définies dans la littérature scientifique, les concepts de résonance ou de prosodie sont plus larges et laissent ainsi place à une plus grande interprétabilité de la part des évaluateurs. De plus, la variabilité de l'évaluation de la qualité vocale dans un contexte de mesure de la parole n'a peut-être pas permis un consensus quant aux critères de jugements sur lesquels chaque auditeur s'est basé. Dans ce contexte, proposer au départ une définition précise des concepts sous-tendus par paramètre spécifique permettrait de limiter ces effets de variabilité. De plus, pour améliorer la pratique clinique, il semble pertinent d'explorer si une

analyse automatique de ces paramètres spécifiques, basée sur des définitions consensuelles de leurs concepts, permet de s'affranchir du biais de variabilité inter-juges (18). Un traitement automatique du signal vocal ciblant des paramètres spécifiques permettrait d'obtenir des scores fiables entre cliniciens ainsi que pour les évaluations de suivi (36,37).

Ainsi, les approches d'évaluation perceptive et automatique de la parole après cancer de la cavité buccale ou de l'oropharynx sont complémentaires. L'automatisation de certaines mesures de bas niveau, comme les altérations phonémiques, les paramètres du signal vocal (dont par exemple la variation de la fréquence fondamentale en jeu dans la prosodie) ou la résonance fournirait des éléments plus précis pour cibler des stratégies thérapeutiques spécifiques. Elle serait malgré tout à compléter par une évaluation perceptive, qui ciblerait la compréhensibilité de la parole, c'est-à-dire la compréhension du message par un auditeur avec la mise en jeu de tous ses processus cognitifs de restauration d'un message altéré.

◆ CONCLUSION

L'influence des différents paramètres spécifiques de parole (qualité vocale, résonance, prosodie et altérations phonémiques) sur la sévérité du trouble diffère selon la localisation tumorale. La qualité vocale n'est en jeu par exemple que pour les patients traités pour un cancer de l'oropharynx en raison de la localisation anatomique de la production de parole. La prosodie n'entre en compte que sur l'ensemble des sujets et disparaît lors des analyses en sous-groupes par localisation. Cet effet peut être lié à un manque de consensus entre les experts quant aux concepts relatifs à la notion de « prosodie » à mesurer dans les productions de parole.

En raison de la variabilité de l'évaluation humaine de ces paramètres spécifiques, une automatisation de l'extraction de ces paramètres permettrait une meilleure fiabilité de l'évaluation, pour adapter les stratégies thérapeutiques au plus près des besoins des patients. Elle viendrait en complément de l'évaluation perceptive classique afin d'avoir des mesures les plus complètes pour structurer les besoins thérapeutiques. De plus, l'étude des stratégies individuelles des auditeurs pour évaluer la parole semble pertinente à analyser. La détermination de profils d'auditeurs-juges permettrait une meilleure compréhension des phénomènes en jeu dans l'évaluation quantitative de la parole, et cibler au mieux l'analyse des résultats perceptifs. Ces derniers resteront indispensables à l'évaluation de la compréhensibilité de la parole, concept fondamentalement humain en raison des mécanismes de compensation cognitifs de restauration d'un message dégradé.

Toutefois, dans une approche holistique, la mesure de l'impact du trouble sur leur communication reste à explorer, en mettant en lien des paramètres issus de l'analyse de différents niveaux linguistiques (voix, phonologie, débit, rythme...), extraits d'un traitement automatique du signal de parole, avec des résultats d'auto-questionnaires ciblant les capacités de communication des patients. Cette évaluation holistique permettrait ainsi de mieux prendre en compte les répercussions fonctionnelles et psychosociales dans le suivi des patients.

◆ CONFLITS D'INTERETS

Aucun conflit d'intérêts n'est déclaré par les auteurs.

◆ REFERENCES BIBLIOGRAPHIQUES

1. Colangelo LA, Logemann JA, Rademaker AW. Tumor size and pretreatment speech and swallowing in patients with resectable tumors. *Otolaryngol Neck Surg.* 2000;122(5):653–61.
2. Dwivedi RC, Kazi RA, Agrawal N, Nutting CM, Clarke PM, Kerawala CJ, et al. Evaluation of speech outcomes following treatment of oral and oropharyngeal cancers. *Cancer Treat Rev* [Internet]. 2009;35(5):417–24. Available from: <http://dx.doi.org/10.1016/j.ctrv.2009.04.013>
3. Borggreven PA, Verdonck-De Leeuw IM, Muller MJ, Heiligers MLCH, De Bree R, Aaronson NK, et al. Quality of life and functional status in patients with cancer of the oral cavity and oropharynx: Pretreatment values of a prospective study. *Eur Arch Oto-Rhino-Laryngology.* 2007;264(6):651–7.
4. DeNittis AS, Machtay M, Rosenthal DI, Sanfilippo NJ, Lee JH, Goldfeder S, et al. Advanced oropharyngeal carcinoma treated with surgery and radiotherapy: Oncologic outcome and functional assessment. *Am J Otolaryngol - Head Neck Med Surg.* 2001;22(5):329–35.
5. Stelzle F, Knipfer C, Schuster M, Bocklet T, Nöth E, Adler W, et al. Factors influencing relative speech intelligibility in patients with oral squamous cell carcinoma: A prospective study using automatic, computer-based speech analysis. *Int J Oral Maxillofac Surg.* 2013;42(11):1377–84.
6. Barrett WL, Gluckman JL, Wilson KM, Gleich LL. A comparison of treatments of squamous cell carcinoma of the base of tongue : Surgical resection combined with external radiation therapy, external radiation therapy alone, and external radiation therapy combined with interstitial radiation. *Brachytherapy.* 2004;3:240–5.
7. Mlynarek A, Rieger J, Harris J, O’Connell D, Al-Qahtani K, Ansari K, et al. Methods of functional outcomes assessment following treatment of oral and oropharyngeal cancer: review of the literature. *J Otolaryngol Head Neck Surg.* 2008;37(1):2–10.
8. Rinkel RN, Leeuw IMV, Van-Reij EJ, Aaronson NK, Leemans R. Speech Handicap Index in patients with oral and pharyngeal cancer: better understanding of patients’ complaints. *Head Neck.* 2008;30:868–74.
9. Reich M. Cancer et image du corps : identité, représentation et symbolique : Le corps retrouvé (French). *Cancer body image identity, Represent.* 2009;85(3):247–54.
10. Middag C, Clapham R, Van Son R, Martens JP. Robust automatic intelligibility assessment techniques evaluated on speakers treated for head and neck cancer. *Comput Speech Lang* [Internet]. 2014;28(2):467–82. Available from: <http://dx.doi.org/10.1016/j.csl.2012.10.007>
11. Pommée T, Balaguer M, Mauclair J, Pinquier J, Woisard V. Assessment of adult speech disorders: current situation and needs in French-speaking clinical practice. *Logop Phoniatr Vocology* [Internet]. 2021;0(0):1–15. Available from: <https://doi.org/10.1080/14015439.2020.1870245>
12. Kent RD, Weismer G, Kent JF, Rosenbek JC. Toward Phonetic Intelligibility Testing in Dysarthria. *J Speech Hear Disord* [Internet]. 1989;54(4):482. Available from: <http://jshd.pubs.asha.org/article.aspx?doi=10.1044/jshd.5404.482>
13. Carmichael J. (EN)[J. Carmichael] Introducing Objective Acoustic Metrics for the Frenchay Dysarthria Assessment Procedure. *Univ Sheff.* 2007;(December).
14. Hustad KC. A closer look at transcription intelligibility for speakers with dysarthria: Evaluation of scoring paradigms and linguistic errors made by listeners. *Am J Speech-Language Pathol.* 2006;15(3):268–77.
15. Hustad KC. The Relationship Between Listener Comprehension and Intelligibility Scores for Speakers With Dysarthria. 2008;51(June 2008):562–73.

16. Balaguer M, Boisguérin A, Galtier A, Gaillard N, Puech M, Woisard V. Assessment of impairment of intelligibility and of speech signal after oral cavity and oropharynx cancer. *Eur Ann Otorhinolaryngol Head Neck Dis* [Internet]. 2019;136(5):355–9. Available from: <https://doi.org/10.1016/j.anorl.2019.05.012>
17. Middag C. Automatische analyse van pathologische spraak Automatic Analysis of Pathological Speech. 2013.
18. Fex S. Perceptual evaluation. *J Voice*. 1992;6(2):155–8.
19. Van Nuffelen G, Middag C, De Bodt M, Martens J. Speech technologybased assessment of phoneme intelligibility in dysarthria. *Int J Lang Commun Disord*. 2009;44(5):716–30.
20. Balaguer M, Pommée T, Farinas J, Pinquier J, Woisard V, Speyer R. Effects of oral and oropharyngeal cancer on speech intelligibility using acoustic analysis: Systematic review. *Head Neck*. 2019;(August):111–30.
21. Astésano C, Balaguer M, Farinas J, Fredouille C, Gaillard P, Ghio A, et al. Carcinologic Speech Severity Index Project: A Database of Speech Disorder Productions to Assess Quality of Life Related to Speech After Cancer. *Lang Resour Eval Conf*. 2018;
22. Meyer TK, Kuhn JC, Campbell BH, Marbella AM, Myers KB, Layde PM. Speech Intelligibility and Quality of Life in Head and Neck Cancer Survivors. *Laryngoscope* [Internet]. 2004;114(11):1977–81. Available from: <http://doi.wiley.com/10.1097/01.mlg.0000147932.36885.9e>
23. Pauloski BR, Logemann JA, Rademaker AW, McConnel FMS, Stein D, Beery Q, et al. Speech and swallowing function after oral and oropharyngeal resections: One-year follow-up. *Head Neck*. 1994;16(4):313–22.
24. Lindblom B. On the Communication Process: Speaker-Listener Interaction and the Development of Speech. *Augment Altern Commun*. 1990;6(4):220–30.
25. Keintz CK, Bunton K, Hoit JD. Influence of visual information on the intelligibility of dysarthric speech. *Am J Speech Lang Pathol* [Internet]. 2007;16(3):222–34. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/17666548>
26. Fontan L. De la mesure de l'intelligibilité à l'évaluation de la compréhension de la parole pathologique en situation de communication. Université Toulouse 2 Le Mirail; 2012.
27. Ghio A, Pouchoulin G, Teston B, Pinto S, Fredouille C, De Looze C, et al. How to manage sound, physiological and clinical data of 2500 dysphonic and dysarthric speakers? *Speech Commun*. 2012;54(5):664–79.
28. Yorkston KM, Strand EA, Kennedy MRT. Comprehensibility of Dysarthric Speech: Implications for Assessment and Treatment Planning. *Am J Speech-Language Pathol*. 1996;5(1):55–65.
29. Woisard V, Astésano C, Balaguer M, Farinas J, Fredouille C, Gaillard P, et al. C2SI corpus: a database of speech disorder productions to assess intelligibility and quality of life in head and neck cancers. *Lang Resour Eval*. 2020;
30. Cicchetti D V. Guidelines, Criteria, and Rules of Thumb for Evaluating Normed and Standardized Assessment Instruments in Psychology. *Psychol Assess*. 1994;6(4):284–90.
31. Mukaka MM. Statistics corner: A guide to appropriate use of correlation coefficient in medical research. *Malawi Med J*. 2012;24(3):69–71.
32. Stelck EH, Boliek CA, Hagler PH, Rieger JM. Current practices for evaluation of resonance disorders in North America. *Semin Speech Lang*. 2011;32(1):58–68.

33. Vaissière J. Language-independent prosodic features. In: Springer Verlag, editor. *Prosody: Models and Measurements*. 1983. p. 53–65.
34. Saravanan G, Ranganathan V, Gandhi A, Jaya V. Speech outcome in oral cancer patients - Pre- and post-operative evaluation: A cross-sectional study. *Indian J Palliat Care*. 2016;22(4):499–503.
35. Schuster M, Stelzle F. Outcome measurements after oral cancer treatment : speech and speech-related aspects — an overview. *Oral Maxillofac Surg*. 2012;291–8.
36. Laaridh I, Fredouille C, Ghio A, Lalain M, Woisard V. Automatic prediction of speech intelligibility based on X-vectors in the context of head and neck cancer. *Proc Annu Conf Int Speech Commun Assoc INTERSPEECH*. 2018;2018–1266(September).
37. Quintas S, Mauclair J, Woisard V, Pinquier J. Automatic prediction of speech intelligibility based on X-vectors in the context of head and neck cancer. *Proc Annu Conf Int Speech Commun Assoc INTERSPEECH*. 2020;2020-October:4976–80.