Evaluation de la compréhensibilité et conservation des fonctions prosodiques en perception de la parole de patients post traitement de cancers de la cavité buccale et du pharynx

Olivier Nocaudie, Corine Astésano, Alain Ghio, Muriel Lalain, Virginie Woisard

▶ **To cite this version:**

Olivier Nocaudie, Corine Astésano, Alain Ghio, Muriel Lalain, Virginie Woisard. Evaluation de la compréhensibilité et conservation des fonctions prosodiques en perception de la parole de patients post traitement de cancers de la cavité buccale et du pharynx. XXXIIe Journées d'Etudes sur la Parole, Jun 2018, Aix-en-Provence, France. pp.196-204, 10.21437/jep.2018-23 . hal-01962272

HAL Id: hal-01962272
https://hal.science/hal-01962272

Submitted on 9 Jan 2019

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.

Evaluation de la compréhensibilité et conservation des fonctions prosodiques en perception de la parole de patients post traitement de cancers de la cavité buccale et du pharynx

Olivier Nocaudie[1] Corine Astésano[1,] Alain Ghio[2], Muriel Lalain[2], Virginie Woisard[1, 3]

(1) Octogone-Lordat (E.A. 4156), Université de Toulouse, France
(2) Aix-Marseille Univ, CNRS, LPL, UMR 7309, Aix-en-Provence, France
(3) C.H.U. de Toulouse-Rangueil, France
nocaudie@univ-tlse2.fr

RESUME

Dans le cadre du projet "Carcinologic Speech Severity Index", cette étude propose l'évaluation perceptive de la parole de 45 locuteurs (37 patients ayant subi des traitements pour des cancers de la cavité buccale et/ou du pharynx, 8 témoins). 66 auditeurs naïfs ont effectué 4 tâches perceptives investiguant, d'une part, la conservation de fonctions prosodiques (groupements syntaxiques, focus pragmatique, modalisation d'énoncés) chez ces patients et, d'autre part, leur compréhensibilité (paradigme de Sentence Verification). Les résultats indiquent une conservation globale des fonctions prosodiques testées chez les patients. En parallèle, ces résultats montrent que le paradigme de Sentence Verification permet de discriminer les locuteurs en fonction de leur compréhensibilité perçue et que cette tâche présente une corrélation forte avec l'évaluation subjective de la sévérité de la parole par des praticiens hospitaliers. Ces résultats sont discutés en regard du rôle joué par la prosodie dans l'adaptabilité communicative des locuteurs en cas de déficience articulatoire.

ABSTRACT

Comprehensibility evaluation and conservation of prosodic function in the speech of patients after Head and Neck carcinologic treatment

In the framework of the Carcinologic Speech Severity Index, the present study proposes to perceptually evaluate speech of 45 speakers (37 patients, with treatment for Head and Neck cancer, 8 controls). 66 naïve listeners fulfilled 4 tasks, investigating on the one hand the conservation of 3 prosodic functions in these patients' speech (syntactic grouping, pragmatic focus, expression of modality) and their intelligibility/comprehensibility using a Sentence Verification Task on the other hand. Results indicate a global conservation of the tested prosodic functions in these patients' speech. The Sentence Verification paradigm results however help discriminate groups of speakers based on their perceived comprehensibility and show higher correlation with subjective evaluation of speech severity by medical practitioners. The results are discussed with regards to the communication adaptative role of prosody in the case of articulatory impairments.

MOTS-CLES : Parole carcinologique, perception, compréhensibilité, prosodie

KEYWORDS: Carcinologic speech, perception, intelligibility, prosody

10.21437/JEP.2018-23

1 Introduction

Les traitements des cancers des voies aérodigestives supérieures (VADS), du fait de leur caractère souvent mutilant malgré les reconstructions anatomiques, peuvent avoir un impact important sur les fonctions de communication. En effet, les organes atteints par ces pathologies sont impliqués dans la production de la parole, notamment au niveau articulatoire (cavité buccale, langue, pharynx). En conséquence, les traitements de ces cancers (i.e. chirurgie, et/ou chimiothérapie) perturbent la production de la parole des patients et, corollairement, sa réception par leurs interlocuteurs.

En raison de la diminution de la mortalité liée aux traitements des cancers, l'évaluation des résultats fonctionnels du traitement carcinologique est une préoccupation majeure des équipes soignantes pour envisager un pronostic fonctionnel concernant la qualité de vie à venir du patient, pour ajuster les protocoles thérapeutiques et améliorer la vie communicationnelle des patients.

1.1 Carcinologic Speech Severity Index (C2SI)

Le projet Carcinologic Speech Severity Index (ci-après C2SI) vise à établir un indice automatique mesurant l'impact des traitements carcinologiques des VADS (*i.e*, ici, des cancers localisés dans la cavité buccale et/ou le pharynx) sur la qualité de la communication des patients. A cette fin, les acteurs du projet C2SI étudient un même corpus original, composé de nombreuses tâches linguistiques (tenues de voyelles, lecture de texte, production de pseudo-mots, description d'images, etc.) recueilli auprès de 135 sujets (Astésano *et al.*, 2018), que ce soit sur le plan acoustique, (Sicard *et al.*, 2017), en perception de la parole (Ghio *et al*., 2017 ; Nocaudie *et al.*, 2017) ou bien en traitement automatique (Lahrid *et al*., 2017).

Pour mettre au point le C2SI, un index automatique évaluant la dégradation de la communication parlée post cancer, il nous est nécessaire d'envisager la parole de ces patients du point de vue de sa réception par les auditeurs humains. D'une part, le recours à l'évaluation perceptive de la parole permet d'obtenir des points de comparaison (Hermes, 1998) pour estimer la validité perceptive d'un indice automatique, *e.g.* C2SI, même si l'holisme de ce procédé présente certaines limites. D'autre part, le handicap lié aux séquelles des traitements des cancers des VADS concerne particulièrement la réception de cette parole altérée. En effet, les difficultés d'articulation dues à la modification des organes de la parole conduiraient à une perte d'intelligibilité du message, *i.e.* de l'aisance des auditeurs à récupérer le détail phonétique du message et parallèlement à une altération de la compréhensibilité du message, *i.e.* la capacité de l'auditeur à interpréter le message indépendamment de ses altérations (Woisard et al..2013). Dans ces termes, l'intelligibilité renvoie au traitement des informations de bas niveaux, tandis que la compréhensibilité évoque des stratégies de haut niveau : par exemple, reconstruction du sens, compensations de l'auditeur ou jugements de plausibilité à partir des connaissances préalables.

1.2 Prosodie et compréhensibilité

En parole, les indices prosodiques se trouvent à l'interface des autres niveaux linguistiques et remplissent de nombreuses fonctions participant à l'encodage comme au décodage des messages. La prosodie permet notamment de marquer la structuration des énoncés en indiquant les frontières des unités, remplissant alors une fonction syntaxique, de moduler la modalité d'un même énoncé par des variations d'indices prosodiques (forme du contour, intensité, etc.) ou encore de faire ressortir l'information centrale d'un message par son marquage accentuel (par exemple,

production d'accent initial emphatique). La prosodie joue alors un rôle pragmatique et informationnel (Di Cristo, 2012).

Au-delà de ces fonctions communicatives, la production cohérente des indices prosodiques d'une langue participe à la fluence des énoncés et à la gestion des aspects temporels de la parole. Par ses fonctions, et ses interactions avec les autres niveaux linguistiques, la prosodie est donc essentielle à la compréhensibilité des messages.

Certains modèles voient d'ailleurs la prosodie comme faisant partie de stratégies compensatoires/palliatives à des altérations segmentales, *i.e.* les locuteurs tendraient à amplifier les indices prosodiques dans leurs productions pour compenser une perte d'intelligibilité/compréhensibilité, due au bruit ambiant ou à des difficultés de prononciation (théorie de l'adaptabilité de la parole à des fins d'optimisation de la communication ; Lindblom, 1990).

1.3 Cadre de l'étude

Dans le cadre du projet C2SI, cette étude propose ainsi d'évaluer d'une part, la compréhensibilité des locuteurs au moyen d'une tache de compréhension de phrases (*Sentence Verification Task* ou SVT, d'après Pisoni & Dedina, 1986) et d'autre part le degré de conservation de trois fonctions prosodiques à l'interface avec la syntaxe, l'expression des modalités et la pragmatique. En effet, malgré une atteinte articulatoire consécutive au(x) traitement(s) dont ils ont bénéficié, les patients parviennent-ils à produire de manière satisfaisante les indices prosodiques, *i.e.* les fonctions prosodiques sont-elles conservées chez ces locuteurs, malgré une perte d'intelligibilité/compréhensibilité souvent constatée par les cliniciens ?

1.4 Hypothèses

Théoriquement, les traitements de ces cancers des VADS ne sont pas supposés impacter la production du flux laryngé et des indices prosodiques, même si certains types de traitement peuvent conduire à une rigidification des tissus au-delà de la zone traitée et, conséquemment à une altération fonctionnelle du larynx. Ainsi, nous posons l'hypothèse que les locuteurs du C2SI obtiendront des résultats satisfaisants aux évaluations perceptives pour ce qui concerne la conservation des fonctions prosodiques malgré ces risques périphériques : il sera difficile de distinguer les groupes patients/témoins sur la seule base de leurs scores, notamment durant les tâches où l'information segmentale est négligeable (syntaxe & modalité). Cependant, nous pensons que l'atteinte articulatoire des patients aura un impact sur le temps de réponse des auditeurs en perception. En effet, l'altération de la parole des patients se traduirait par des difficultés de traitement par les auditeurs. Il est donc en revanche hautement probable que le groupe patient obtienne des résultats moins bons que ceux du groupe contrôle au test de compréhensibilité (SVT), en raison de ces difficultés de traitement de la parole altérée par les auditeurs.

2 Aspects méthodologiques

2.1 Recueil du matériau linguistique

Une cohorte de 135 locuteurs enregistrés dans le service d'oncoréhabilitation de l'Oncopole à Toulouse (94 patients / 41 sujets contrôles) a participé à un recueil de données de parole constitué de l'ensemble des tâches de C2SI (Astésano *et al*. 2018). Pour cette étude, nous avons sélectionné

une sous-partie du corpus C2SI comprenant les productions de 45 locuteurs (37 sujets patients et 8 sujets contrôles). Ces productions correspondent à la première phase d'enregistrement de C2SI.

Notre étude porte donc sur quatre tâches de C2SI, trois évaluant de grandes fonctions prosodiques (groupements syntaxiques, marquage du focus pragmatique, marquage de la modalité) et une destinée à l'évaluation de la compréhensibilité (Sentence Verification Task).

2.1.1 Prosodie : Groupement syntaxique

Les locuteurs devaient résoudre une ambiguïté syntaxique par des moyens prosodiques dans des syntagmes composés de deux noms et un adjectif (Astésano, Bard & Turk (2007). L'ambiguïté réside dans la portée de l'adjectif qui peut alternativement qualifier le second nom du syntagme, ou bien les deux noms :

a. [Les tomates][et les **oignons rouges**]

b. [Les **tomates** et les **oignons**][**rouges**]

Les énoncés étaient présentés en modalité écrite et le groupement syntaxique attendu marqué par des indications visuelles. Chaque locuteur a enregistré le même ensemble de 26 énoncés (13 scripts différents en ordre aléatoire, par bloc de groupement syntaxique * 2 conditions syntaxiques).

2.1.2 Prosodie : Marquage du Focus pragmatique

Cette tâche demandait aux locuteurs de marquer le focus pragmatique en mettant en valeur l'information importante d'un énoncé par des moyens prosodiques (Astésano *et al*., 2004; Magne *et al*, 2005. Corpus adapté pour les patients par Aura, 2012). Suite à l'écoute d'une question indiquant quel focus produire « *Où as-tu vu un canard, dans le jardin ou dans la cour* ? » *vs.* « *Qu'as-tu vu dans le jardin, un cochon ou un canard* ? », les sujets devaient lire la réponse en marquant le focus contrastif de manière adéquate : (« *j'ai vu un canard **dans le jardin*** » *vs.* (« *j'ai vu un **canard** dans le jardin* »). Chaque locuteur a enregistré le même ensemble de 20 énoncés (10 contenus lexicaux * 2 focus pragmatiques, présentés selon un ordre aléatoire).

2.1.3 Prosodie : Expression de la Modalité

Lors de cette tâche les locuteurs devaient produire des énoncés identiques au niveau lexical, *e.g.* « *tu prends la voiture* », en les modulant au niveau prosodique, afin de transmettre trois modalités différentes (assertion, interrogation, ordre). Les énoncés étaient présentés en modalité écrite, et la modulation prosodique demandée était marquée par un symbole de ponctuation ('.' ; ' ?' ; ' !'). Chaque locuteur a enregistré un même ensemble de 30 énoncés (10 scripts en ordre aléatoire *3 blocs de modalité).

2.1.4 Compréhensibilité : Sentence Verification Task

Une liste de 150 couples de phrases vraies/fausses et dont le caractère vrai ou faux ne peut être vérifié qu'à partir de la dernière unité lexicale a été constituée (*e.g.* « *La poule pond des œufs* » vs. « *La poule pond des fruits* »). Certains énoncés sont adaptés des propositions de Pisoni & Dedina (1986), d'autres de celles de Zumbiehl (2010), les derniers couples de phrases ont été constitués par les membres du projet C2SI. Chaque locuteur a enregistré un ensemble de 50 énoncés tirés des listes SVT.

Notre corpus pour cette étude représente donc un ensemble de 5670 énoncés [45 sujets * (26 items syntaxe + 30 items modalité + 20 items focus + 50 items SVT)]

2.2 Tâches en perception

Suite à ce recueil, une cohorte de 66 auditeurs (18-31 ans, 54 femmes/12 hommes) naïfs (*i.e.* non habitués à la parole altérée pour éviter l'effet des processus de haut niveau dans la perception de la parole, Ohala, 1986) ont procédé à l'évaluation perceptive de ces trois fonctions prosodiques ainsi que de la compréhensibilité des locuteurs. Les auditeurs étaient indemnisés (10 €, une heure de passation) pour leur participation à l'expérience qui s'est déroulée sur la plateforme « Comportement Cognition Usages » de l'université Toulouse–Jean Jaurès.

L'ensemble des tâches a été scripté pour le logiciel PERCEVAL (André et al., 2003), de manière à ce que chaque énoncé produit lors du recueil soit évalué par au moins trois auditeurs différents. Chaque tâche décrite lors du recueil équivalait à un bloc expérimental en perception, au sein duquel les items étaient randomisés. Les blocs étaient eux-mêmes présentés dans un ordre aléatoire. Enfin, notons qu'avant chaque bloc, une phase d'entraînement était accomplie par les auditeurs pour se familiariser avec les tâches.

2.2.1 Tâche de perception des frontières syntaxiques (SYN)

Lors de la tâche SYN, il était demandé aux participants d'indiquer la structure syntaxique comprise après écoute d'un stimulus (ex : [Les tomates][et les **oignons rouges**] vs [Les **tomates** et les **oignons**][**rouges**]). Une image rappelant les cas possibles était présentée à l'écran lors de l'écoute. Chaque participant évaluait 50 à 55 énoncés (en fonction du groupe).

2.2.2 Tâche de perception du focus pragmatique (FOC)

Pour construire cette tâche de perception du focus pragmatique, nous avons associé des questions impliquant la production du focus pragmatique (les mêmes enregistrements que ceux utilisés lors du recueil du corpus) avec des réponses enregistrées par les patients. Nous avons ainsi créé des dialogues question/réponse en manipulant la congruité/incongruité du dialogue pour ce qui concerne la production du focus pragmatique : « *Où as-tu vu un canard, dans le jardin ou dans la cour* ? », puis « *j'ai vu un CANARD dans le jardin* » (réponse inappropriée) *vs.* « *j'ai vu un canard dans le JARDIN* » (réponse appropriée). Après écoute de chaque dialogue, les participants devaient indiquer si la réponse à la question était appropriée ou non. Chaque participant évaluait un ensemble de *ca*.40 dialogues.

2.2.3 Tâche de perception de la modalité (MOD)

Durant le bloc MOD, les participants écoutaient un énoncé avant de devoir en indiquer la modalité comprise parmi trois possibles (interrogation, assertion, ordre). Chaque participant évaluait *ca.* 61 énoncés.

2.2.4 Sentence Verification Task (SVT)

Pour la SVT, les énoncés vrai/faux recueillis lors des enregistrements étaient présentés auditivement aux participants qui devaient indiquer si la phrase entendue était plutôt vraie (« *la poule pond des œufs* ») ou plutôt fausse (« *la poule pond des fruits* »). Chaque participant devait vérifier la véracité de *ca.* 100 phrases.

Pour chacune de ces tâches, nous avons calculé un score par locuteur qui équivaut à la moyenne de chacune des évaluations perceptives reçues durant le test. Dans la mesure où chaque item a été évalué trois fois à travers les auditeurs, ce score est compris entre 0 et 3. Par ailleurs, les temps de réaction des auditeurs ont été calculés (du début de l'écoute de l'énoncé à la réponse, temps auquel est soustrait la durée de l'énoncé) et moyennés pour chacun des locuteurs, afin qu'ils soient adjoints à nos résultats. Ainsi, chaque locuteur du C2SI a obtenu 4 scores différents (un par tâche) associés à des temps de réaction moyens des auditeurs.

3 Résultats

3.1 Différences intergroupes : Témoins *vs.* Patients

Dans la TABLE 1, nous indiquons pour chacun de nos groupes (Témoins vs. Patients) des indicateurs de statistiques descriptives afin d'observer des tendances de performances d'une part et de variations intragroupes d'autre part.

Tâche	**Variable**	**Témoins**			**Patients**		
		Moyenne	Médiane	Ecart type	Moyenne	Médiane	Ecart type
SVT	Scores	2,91	2,92	0,04	2,36	2,75	0,66
	TR	1049	1046	119	2750	1699	2078
FOC	Scores	2,59	2,66	0,21	2,33	2,40	0,35
	TR	987	965	413	1891	1820	796
SYN	Scores	2,21	2,22	0,39	1,81	1,69	0,39
	TR	2418	2363	277	2428	2219	563
MOD	Scores	2,17	2,37	0,53	1,77	1,79	0,44
	TR	1754	1654	278	2262	2256	424

TABLE 1 : Indices de variabilité individuelle par groupes. Les scores sont compris entre 0 et 3. Les temps de réaction (TR) sont donnés en millisecondes.

Ainsi, le groupe « Témoins » présente des scores moyens plus élevés ainsi que des temps de réaction moyens plus court que le groupe « Patient ». Cependant, la différence la plus flagrante entre les deux groupes est illustrée par les écarts types du groupe « Patient », beaucoup plus élevés, qui témoignent d'une grande dispersion des données au sein du groupe (tant dans les scores moyens reçus que dans les temps moyens de réaction relevés), notamment pour ce qui concerne la tâche de SVT. Ceci étant dit, les scores moyens obtenus par les deux groupes aux tâches de modalité et de syntaxe sont décevants, plus particulièrement pour le groupe Témoin.
Afin de confirmer ces différences intergroupes en termes de scores moyens et temps de réaction moyens, nous avons conduit des Anovas à un facteur (Témoins vs Patient). Les résultats de ces tests indiquent des différences statistiquement significatives en ce qui concerne les **scores reçus** à la tâche de focus ($F(1, 773) = 12.72$, $p.<.001$), de jugement de la modalité ($F(1, 1137) = 22.50$, $p.<001$), de groupement syntaxique ($F(1, 1038) = 25.43$, $p.<.001$) ainsi qu'à la SVT ($F(1, 1825) = 71.25$, $p.<.001$). A propos des **temps de réaction**, nos tests ont indiqué des différences statistiquement significatives pour toutes les tâches [focus ($F(1, 773) = 52.84$, $p.<.001$) ; modalité ($F(1, 1137) = 37.57$, $p.<001$) ; SVT ($F(1, 1825) = 65.81$, $p.<.001$)], excepté pour la tâche de groupement syntaxique ($F(1, 1038) = 0.007$, *ns*).

3.2 Corrélation entre résultats des tests perceptifs et évaluations cliniques

Finalement, nous avons voulu savoir si les évaluations accordées par notre jury d'auditeurs naïfs reflétaient les résultats obtenus par les locuteurs par rapport aux évaluations cliniques. En effet, chaque locuteur a été évalué par 5 professionnels de santé sur deux indices : un indice de sévérité et un indice d'intelligibilité. Dans la mesure où les résultats de nos corrélations étaient très proches pour les deux indices, nous ne présentons dans la TABLE 2 que les corrélations avec l'indice de sévérité.

Tâche	Facteur	r	p value
SVT	Score moyen	0.81	<.001
	Temps de réaction	-0.76	<.001
FOC	Score moyen	0.56	<.001
	Temps de réaction	-0.67	<.001
MOD	Score moyen	0.44	<.05
	Temps de réaction	-0.66	<.001
SYN	Score moyen	0.53	<.001
	Temps de réaction	-0.29	ns

TABLE 2 : résultats des tests de corrélation de Pearson entre l'indice de sévérité et le score moyen obtenu par les locuteurs/le temps de réaction des auditeurs. Les valeurs de l'indice de sévérité sont comprises entre 0 (altération très sévère) et 10 (aucune altération).

Ainsi, la corrélation entre score moyen à la SVT et sévérité estimée du locuteur révèle un lien positif fort entre les deux variables ($r = 0.81$, $p.<.001$). Les tests de corrélation avec les scores obtenus aux autres tâches montrent également des liens positifs, mais plus mesurés, s'échelonnant de $r = 0.44$ (MOD) pour le plus faible à $r = 0.56$ (FOC) pour la plus modérément forte.
Par conséquent, la SVT semble constituer une tâche pertinente pour estimer la compréhensibilité du locuteur.

A propos des corrélations entre temps de réaction moyen des auditeurs et indice de sévérité, comme attendu, chaque test indique une relation négative modérément forte à forte [$-0.76 < r < -0.66$] pour l'ensemble des tâches, indiquant alors que plus l'indice de sévérité estimé est bas (*i.e.* plus l'atteinte du locuteur est sévère) plus le temps de réaction de l'auditeur lors de la tâche est élevé, exception faite de la tâche de syntaxe, où le score de corrélation est faible (et non significatif).

4 Discussion & Conclusion

Nos résultats présentent plusieurs tendances différentes en fonction des tâches effectuées en perception.
Les résultats de la SVT sont les plus nets, les Patients ayant une moins bonne performance que les Témoins et provoquant des temps de réaction plus longs chez les auditeurs. Cela dit, le groupe Patient montre, particulièrement pour cette tâche, une grande variabilité, les meilleurs atteignant une performance proche de certains Témoins. De plus, les résultats de cette tâche présentaient une corrélation forte avec l'évaluation perceptive de la sévérité des locuteurs par des praticiens hospitaliers. La SVT constituerait ainsi une tâche adéquate pour évaluer la compréhensibilité des locuteurs. Cependant, une étude plus poussée des différences interindividuelles (base de données cliniques C2SI) au sein du groupe patient nous semble nécessaire pour expliquer la variabilité observée (*cf.* TABLE 1).
En prosodie, les résultats étaient moins tranchés. La tâche de focus, a présenté la meilleure corrélation avec l'évaluation clinique, les meilleurs scores moyens et les meilleurs temps de

réaction pour les deux groupes, bien que les patients aient été significativement moins performants que les Témoins. Pour affiner ces résultats, il conviendrait de passer sur le plan acoustique pour observer la réalisation des indices acoustique des accents initiaux emphatiques ainsi que leur localisation dans l'énoncé.

A propos de la tâche de syntaxe et, dans une moindre mesure, de la tâche de modalité, les témoins se sont peu distingués des patients (différences plus réduites que dans les autres tâches, voire non significatives). De plus, les scores moyens obtenus se rapprochent des seuils de hasard. Ceci laisse penser que les locuteurs enregistrés ont pu avoir des difficultés à produire les effets attendus, par exemple, en raison de la difficulté de la tâche de groupement syntaxique qui reste peu naturelle, y compris pour les locuteurs sains (voir Astésano *et al.*, 2007, pour une discussion). En ce qui concerne la production des modalités, la question se pose de savoir si les contours intonatifs produits par les locuteurs étaient suffisamment prototypiques des trois modalités (question, ordre, assertion) pour être reconnus comme tels par les auditeurs. Ces questions constituent des perspectives à développer en prolongement de ce travail, en utilisant des métriques de comparaison des formes prosodiques (Nocaudie, 2016) pour s'assurer que la réalisation phonétique des groupements syntaxiques ou des modalités sont prototypiques.

Finalement, l'autre tendance nette de nos résultats est constituée par la différence de temps moyens de réaction obtenus qui sont systématiquement plus longs pour le groupe patient. Cela pourrait révéler des difficultés de traitement de l'information chez nos auditeurs naïfs, qui n'ont pas l'habitude d'être exposés à de la parole altérée et n'avaient pas la possibilité de réécouter plusieurs fois les stimuli. En SVT, nous pourrions admettre que les locuteurs étaient parfois incapables de répondre car ils n'arrivaient pas à récupérer l'information lexicale. En prosodie, cela semble s'échelonner en fonction des tâches. La tâche FOC ressemble à la SVT (et suit les mêmes tendances), mais dans le domaine prosodique : le locuteur doit maintenir une compréhensibilité suffisante des unités lexicales sur l'énoncé entier et produire les indices prosodiques pertinents. En tâche de MOD ou SYN, les temps de réaction des Témoins et des Patients tendent à se rejoindre : le contour prosodique pourrait suffire seul à accomplir la tâche mais des altérations articulatoires pourraient perturber la perception des contours par les auditeurs (perte de fluence, par exemple). Ainsi, la tâche de focus pourrait représenter une tâche cumulant évaluation de la compréhensibilité et des aspects prosodiques de la parole : évaluation d'indices locaux (réalisation d'IA) et globaux (mesures de fluence, débit, intensité, étendue, etc).

Ce travail doit se poursuivre dans plusieurs directions : en intégrant l'intégralité des locuteurs du corpus C2SI et en exploitant les métadonnées associées (base clinique) dans nos modèles statistiques ainsi que sur le plan acoustique, pour pouvoir décider s'il faut inclure/exclure certaines tâches utilisées dans un outil automatisé type C2SI.

Enfin, cette étude souligne, parfois encore en pointillés, des potentiels en termes de production d'événements prosodiques dans le groupe patient. Si les analyses acoustiques montraient effectivement une bonne conservation des fonctions communicatives et structurels de la prosodie, alors faudrait-il envisager de les exploiter, chez des patients peu intelligibles (indices prosodiques de rythme) et sensibiliser l'entourage des patients à ces enjeux.

Remerciements

Financement n°2014-135 de l'Institut National pour le CAncer (INCA) de novembre 2014, "Sciences Humaines et Sociales, Épidémiologie et Santé Publique". Porteur : Pr Virginie Woisard.

Références

ANDRE C., GHIO A., CAVE C., & TESTON B. (2003). PERCEVAL: a Computer-Driven System for Experimentation on Auditory and Visual Perception. In *Proceedings of XVth ICPhS* (p.1421‑1424). Barcelone, Espagne

ASTÉSANO C., BARD, E. G. & TURK, A. (2007). Structural influences on initial accent placement in French. *Language and Speech*, 50(3), 423-446.

ASTESANO C., BALAGUER, M., FARINAS, J., FREDOUILLE, C., …, WOISARD, V. (2018). Carcinologic Speech Severity Index Project: A Database of Speech Disorders Productions to Assess Quality of Life Related to Speech After Cancer. *Proceedings of the 11th Language Resources and Evaluation Conference*, papier accepté, Miyazaki, Japon

AURA K. (2012). *Protocole d'évaluation du langage fondé sur le traitement de fonctions prosodiques : étude exploratoire de deux patients atteints de gliomes de bas grade en contexte péri-opératoire*. Doctoral dissertation, Université Toulouse le Mirail-Toulouse II.

DI CRISTO A. (2012). Le pouvoir de la prosodie ou la revanche de Cendrillon. In Baqué L. & Estrada M. *L'Homme Communiquant*, CIPA, pp.95-114, 2012.

GHIO A., LALAIN M., GIUSTI L., ROBERT D., POUCHOULIN G., …, WOISARD V. (2017). Du décodage acoustico-phonétique pour mesurer l'intelligibilité de locuteurs atteints de troubles de production de la parole, *7emes Journées de Phonétique clinique*, Paris, 29-30 june

HERMES D. J. (1998). Measuring the Perceptual Similarity of Pitch Contours. *Journal of Speech, Language and Hearing Research*, *41*, 73–82.

LAARIDH I., KHADER B. W., FREDOUILLE C. & MEUNIER C. (2017). Automatic Prediction of Speech Evaluation Metrics for Dysarthric Speech, *Interspeech'17*, Stockholm, Sweden.

LINDBLOM B. (1990) Explaining Phonetic Variation: A Sketch of the H&H Theory. In: Hardcastle W.J., Marchal A. (eds) *Speech Production and Speech Modelling*, vol 55. Springer, Dordrecht

MAGNE C.; ASTÉSANO C.; LACHERET-DUJOUR A.; MOREL M.; ALTER K. & BESSON M. (2005) On-line processing of "pop-out" words in spoken french dialogues. *Journal of Cognitive Neuroscience, 17 (5),* 740-756.

PISONI D. & DEDINA M. (1986) "Comprehension of Digitally Encoded Natural Speech using a Sentence Verification Task: a first report" in *Research on Speech Perception. Progress Report N°12*, Indiana University

OHALA J. (1986). Phonological evidence for top-down processing in speech perception. In *Invariance and variability in speech processes* (Erlbaum, p. 386‑397). Hillsdale: Perkell & Klatt

NOCAUDIE O (2016). *Imitation et contrôle prosodique dans l'entraînement à la remédiation phonétique : évaluation, mesure et applications pour l'enseignant de langue étrangère*. Doctoral dissertation, Université de Toulouse.

NOCAUDIE O., ASTESANO C. & WOISARD V. (2017). Conservation des fonctions prosodiques post traitement des cancers de la cavité buccale et du pharynx, *7emes Journées de Phonétique clinique*, Paris-29-30 june

SICARD E., MAUCLAIR J. AND WOISARD V. (2017). Etude de paramètres acoustiques des voix de patients traités pour un cancer ORL dans le cadre du projet C2SI, *7emes Journées de Phonétique clinique*, Paris, 29-30 june

WOISARD V., ESPESSER R., GHIO A. & DUEZ D. (2013). De l'intelligibilité à la compréhensibilité de la parole, quelles mesures en pratique clinique ?. *Rev Laryngol. Otol. Rhinol*., 134 (1), 27-33.

ZUMBIEHL O. (2010). Evaluation perceptive des dysphonies par la Sentence Verification Task. Mémoire d'Orthophonie (dir. : Cavé C. & Ghio A.). Université Aix-Marseille.