



**HAL**  
open science

## Actes des 9èmes Journées de Phonétique Clinique

Mathieu Balaguer, Lila Gravelier, Julien Pinquier

► **To cite this version:**

Mathieu Balaguer, Lila Gravelier, Julien Pinquier. Actes des 9èmes Journées de Phonétique Clinique : "Prendre la mesure de la parole". Institut de Recherche en Informatique de Toulouse, 2023, 978-2-917490-35-8. hal-04126591v2

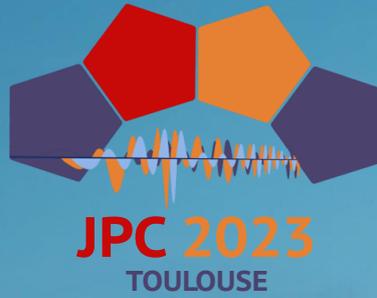
**HAL Id: hal-04126591**

**<https://hal.science/hal-04126591v2>**

Submitted on 28 Jun 2023

**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.



# 9<sup>ème</sup> Journées de Phonétique Clinique

## "Prendre la mesure de la parole"



UNIVERSITÉ TOULOUSE - JEAN JAURÈS  
15-17 JUIN 2023

Conception graphique: © Benoît Colas, UT2J/CPRS-DAR. Photographie: © DR.





# **9<sup>ème</sup> Journée de Phonétique Clinique**

**« Prendre la mesure de la parole »**

Directeurs d'édition : Mathieu Balaguer, Lila Gravelier et Julien Pinquier

Cet ouvrage a été édité et achevé d'imprimer en juin 2023 par  
L'Institut de Recherche en Informatique de Toulouse  
118 Route de Narbonne, F-31062 TOULOUSE CEDEX 9

Dépôt légal : juin 2023



# 9<sup>ème</sup> Journée de Phonétique Clinique

## « Prendre la mesure de la parole »

Voix, parole et surdité

Troubles neurologiques de la parole et de la voix

Prosodie

Evaluation de la parole : vers la clinique

Par Mathieu Balaguer, Lila Gravellier et Julien Pinquier

ISBN : 978-2-917490-35-8

# Table des matières

<b>Avant-Propos</b> .....	<b>v</b>
<b>Comité scientifique</b> .....	<b>vi</b>
<b>Comité d'organisation</b> .....	<b>viii</b>
<b>Communications invitées</b> .....	<b>1</b>
Phénomènes compensatoires : utilité pour l'évaluation et le traitement de la parole déviante.....	3
Co-designing smart technologies for self-expression: case studies with people with intellectual disability .....	5
L'analyse acoustique de la voix en contexte clinique : recommandations actuelles et enjeux cliniques .....	7
<b>Communications orales</b> .....	<b>9</b>
<b>Voix, parole et surdité</b> .....	<b>11</b>
Échelles perceptives pour les voix parlées normophoniques, pathologiques et chantées : revue de littérature et réflexions méthodologiques .....	13
Contribution de l'information segmentale dans l'accès lexical chez des enfants avec implants cochléaires et enfants avec audition typique .....	17
Amélioration de la production de consonnes chez les enfants porteurs d'implants cochléaires lorsqu'ils sont exposés à la Langue française Parlée Complétée : corrélats acoustiques et articulatoires .....	19
Nasalité vocalique, compétences phonologiques et grammaticales : des liens spécifiques chez les enfants porteurs d'implants cochléaires ? .....	23
Analyse Factorielle des Enveloppes de Modulations d'Amplitude des signaux de la parole dans la perspective de l'hypothèse du codage efficace et de l'application aux implants cochléaires...27	
<b>Troubles neurologiques de la parole et de la voix</b> .....	<b>31</b>
Contribution à l'étude psychométrique du Voice Handicap Index auprès de patients parkinsoniens .....	33
Diagnostic différentiel des troubles moteurs de la parole grâce aux mouvements oro-moteurs et aux modulations de la parole .....	35
La désorganisation temporelle dans la dysarthrie ataxique : une altération au niveau global ou local .....	37
Exploration des liens entre fluence, disfluences, organisation temporelle de la parole et performances cognitivo-langagières dans les aphasies primaires progressives .....	39
Prédiction de la fluence à partir de mesures temporelles automatiques : vers une application clinique pour améliorer le diagnostic et le suivi des patients aphasiques .....	43
<b>Prosodie</b> .....	<b>47</b>
Impact de la maladie de Parkinson sur les paramètres acoustiques de la prosodie : une revue systématique et une méta-analyse Bayésienne.....	49

Caractérisation de la prosodie du mot chez des enfants avec autisme avec et sans déficit phonologique .....	51
Variations prosodiques lors de la synchronisation de parole avec un rythme externe : comparaison de locuteurs qui bégaièrent et normo-fluents.....	53
Une nouvelle procédure psychoacoustique pour caractériser la perception de la prosodie après AVC de l'hémisphère droit.....	57
Représentation automatique du rythme de la parole pathologique via le spectre de modulations d'amplitude.....	59
<b>Evaluation de la parole : vers la clinique .....</b>	<b>63</b>
Effet de la tâche sur la voix et la parole de patientes dysphoniques avec suspicion d'atteinte du nerf laryngé supérieur .....	65
Exploration des stratégies compensatoires respiratoires en parole : le cas des fentes palatines.....	68
Contribution de MonPaGe-2.0.s au diagnostic différentiel de troubles moteurs acquis de la parole.....	71
Vers une interprétation, en terme d'altération de traits phonétiques, du niveau d'intelligibilité chez des patients atteints de cancer de la tête ou du cou .....	73
Une approche granulaire pour la prédiction automatique de l'intelligibilité de la parole .....	75
<b>Posters .....</b>	<b>77</b>
<b>Session posters 1.....</b>	<b>79</b>
Utilisation de l'échographie comme outil d'évaluation des troubles des sons de la parole (TSP) dans un contexte clinique québécois : étude pilote sur la connaissance de l'outil et la variabilité des analyses .....	81
Relations entre évaluation acoustique et perceptive de la qualité de la voix : une approche dynamique .....	85
Analyse acoustique de toux volontaires et réflexes : comparaison entre sujets sains et patients avec une tumeur tête et cou.....	89
L'identification des émotions chez des enfants bilingues à partir des indices prosodiques du signal de parole : état de la littérature et proposition méthodologique .....	93
Troubles de la parole dans la maladie de Parkinson : effets d'une prise en charge par l'utilisation de la musique ? Étude de cas.....	97
Fiabilité des valeurs de temps maximal de phonation, débit diadococinétique et débit de parole au travers de multiples essais par des participants neuro-typiques .....	99
Relation entre productions verbales et gestuelles chez deux enfants porteurs d'une trisomie 21 : une étude préliminaire .....	103
Prosodie et gestualité chez les adolescents avec un trouble du spectre autistique .....	105
Modélisation individuelle de juges pour la prédiction automatique de l'intelligibilité de la parole .....	109
La voix de la résilience : suivi acoustique des trajectoires de guérison des patients post-traumatiques au cours des séances de ré-exposition en imagination.....	113

Exploration de caractéristiques linguistiques et acoustiques pour la génération automatique de rapports de séances de remédiation cognitive avec un assistant virtuel.....	117
Reconnaissance automatique des expressions émotionnelles dans la parole naturelle : le cas des patients dans des appels d'urgence à Besançon et à Lausanne.....	119
La polypose naso-sinusienne : impacts spectraux. Étude préliminaire. ....	123
Différents stades d'Alzheimer : Rôle des méthodes automatisées de traitement de la prosodie et de la qualité de la voix.....	127
Etude des signaux vibroacoustiques de déglutition chez les sujets sains.....	129
Paramètres pertinents pour l'évaluation du niveau de lecture : vers l'élaboration d'un outil d'aide à la fluence.....	133
Comparaison de l'effet d'amorces rythmiques vocales, musicales et motrices sur le temps d'initiation de la parole chez des adultes tout-venant.....	135
<b>Session posters 2 .....</b>	<b>141</b>
Glucides raffinés et attractivité vocale .....	143
Élaboration et normalisation d'une tâche évaluant les représentations sémantiques des noms et des verbes chez l'adulte .....	145
La somnolence influence-t-elle la production des schwas chez les patients hypersomniaques ? .....	147
Validation d'un corpus de paires minimales illustrées adapté aux enfants de 7 à 10 ans : Paires minimales Illustrées Francophones (PIF).....	151
Perception de parole chez les enfants porteurs d'implants cochléaires : Apports de l'Auditory Verbal Therapy et de la Langue française Parlée Complétée.....	155
La voix comme marqueur potentiel de l'autisme .....	157
Les patients en état de coma sont-ils sensibles à la prosodie ?.....	159
Flexibilité phonétique et Autisme .....	163
La mesure de l'intelligibilité en lecture par le biais de séquences délexicalisées chez des patients atteints de la maladie de Parkinson.....	167
Comparaison de la structure formantique des voyelles en lecture et en parole spontanée dans la parole bègue et non bègue : une étude pilote.....	169
De la valeur aberrante à l'altération <i>rare</i> de la parole : application à la parole pathologique.....	171
Analyse des performances de systèmes de reconnaissance automatique de la parole spontanée après cancer oral ou oropharyngé.....	175
Diagnostic différentiel chez les enfants avec un trouble des sons de parole : intérêt de l'épreuve de diadococinésies.....	179
Coarticulation anticipatoire V-à-V dans différentes dysarthries.....	181
Similitudes entre les mécanismes de la déglutition et la production de la parole .....	185
Classification d'Alzheimer à partir de paramètres acoustiques et prosodiques avec de l'apprentissage automatique.....	187

Influence d'un métronome sur la fluence et les gestes de production de parole de personnes qui bégaièrent.....	191
Réflexions autour des pauses : quels types de « dysfluencèmes » dans l'aphasie pour quelles mesures en discours continu ?.....	195
<b>Index des auteurs .....</b>	<b>197</b>

# Avant-Propos

Depuis leur création en 2005, les Journées de Phonétique Clinique (JPC) ont été régulièrement organisées sur une base bisannuelle. Après une dernière édition organisée en Belgique par nos collègues du Laboratoire de phonétique de l'Université de Mons en 2019 (sous l'égide de l'Institut de Recherche en Sciences et Technologies du Langage), les JPC reviennent en France en 2023 (après annulation en 2021) pour leur 9<sup>e</sup> édition. Co-organisée par l'Institut de Recherche en Informatique de Toulouse (IRIT), le laboratoire de Neuro-Psycho-Linguistique (LNPL) et le Centre Hospitalo-Universitaire de Toulouse ainsi que par le Laboratoire Informatique d'Avignon (LIA), la manifestation s'est tenue à l'Université Toulouse Jean Jaurès du 15 au 17 juin 2023.

Rencontre scientifique internationale, les Journées de Phonétique Clinique sont principalement destinées à rassembler et à favoriser les échanges entre chercheurs, cliniciens, informaticiens, ingénieurs, phonéticiens et tout autre professionnel s'intéressant au fonctionnement de la parole, de la voix et du langage. Les JPC accueillent autant les experts que les jeunes chercheurs et les étudiants des domaines cliniques (médecine, orthophonie/logopédie), psychologique, informatique et des sciences du langage.

La production et la perception de la parole, de la voix et du langage de l'enfant et de l'adulte, sain ou atteint d'une pathologie, sont les domaines de prédilection des JPC. Ils y sont ainsi abordés selon des points de vue variés, permettant le partage des savoirs et l'ouverture de nouvelles pistes de réflexion, de recherche et de collaboration.

Lors de cette neuvième édition, la thématique des **mesures de la parole** sera mise en avant. Elle s'inscrit dans un cadre conceptuel dont les facettes sont multiples : analyses perceptives, traitement automatique du signal, caractérisations de l'intelligibilité, du trouble de la parole, des di/ysfluences atteignant le débit de la parole, la prosodie... Sa pertinence clinique est essentielle : l'évaluation du trouble, de ses conséquences fonctionnelles et de l'impact sur la qualité de vie est primordiale pour le suivi des patients atteints de pathologies neurologiques, cancérologiques...

Les propositions de communication portent sur (liste non exhaustive) :

- Parole et perturbations des systèmes perceptifs, auditifs et visuels
- Modélisation de la parole et de la voix pathologiques
- Perturbations du système oro-pharyngo-laryngé
- Évaluation fonctionnelle de la parole, du langage et de la voix.
- Diagnostic et traitement des troubles de la parole et de la voix parlée et chantée
- Instrumentation et ressources en phonétique clinique
- Troubles cognitifs et moteurs de la parole et du langage

# Comité scientifique

- Martine Adda-Decker, LPP, Sorbonne Université
- Maxime Amblard, LORIA, Université de Lorraine
- Angélique Amelot, LPP, Université Sorbonne Nouvelle
- Benoît Amy de la Bretèque, PH, CHRU de Montpellier
- Régine André-Obrecht, IRIT, UPS, Université de Toulouse
- Corine Astésano, LNPL, UT2J, Université de Toulouse
- Nicolas Audibert, LPP, Université Sorbonne Nouvelle
- Pierre Badin, GIPSA-lab, Université de Grenoble Alpes
- Mathieu Balaguer, IRIT, UPS, CHU/Université de Toulouse
- Lorraine Baqué, Universitat Autònoma de Barcelona
- Melissa Barkhat-Defradas, ISEM, Université Montpellier 2
- Stéphanie Borel, Sorbonne Université
- Véronique Boulenger-Mozdzen, DDL, Lyon Université
- Lise Crevier-Buchman LPP, Université Paris 3 & Hôpital Georges Pompidou
- Olivier Crouzet, LLING, Université de Nantes
- Véronique Delvaux, FNRS & IRSTL, Université de Mons
- Christelle Dodane, Praxiling, Université Montpellier 3
- Jérôme Farinas, IRIT, UPS, Université de Toulouse
- Camille Fauth, LILPA, Université de Strasbourg
- Sandrine Ferré, Université de Tours
- Lionel Fontan, Archean Labs
- Cécile Fougeron, LPP, Sorbonne Université
- Corinne Fredouille, LIA, Avignon Université
- Alain Ghio, LPL, Aix-Marseille Université
- Bernard Harmegnies, Université de Mons
- Nathalie Henrich Bernardoni, GIPSA-lab, Université de Grenoble
- Fabrice Hirsch, Praxiling, Université Montpellier 3
- Kathy Huet, Laboratoire de Phonétique, Université de Mons
- Marina Laganaro, Université de Genève
- Muriel Lalain, LPL, Aix-Marseille Université
- Yves Laprie, LORIA, Université de Lorraine
- Marie-Thérèse Le Normand, LPPS, Université Paris Descartes
- Thierry Legou, LPL, Aix Marseille Université
- Anna Marczyk-Buklaha, LNPL, UT2J, Université de Toulouse
- Julie Mauclair, IRIT, UPS, Université de Toulouse
- Christine Meunier, LPL, Aix-Marseille Université
- Yohann Meynadier, LPL, Aix-Marseille Université
- Michel Mondain, CHU de Montpellier
- Dominique Morsomme, Université de Liège
- Thomas Pellegrini, IRIT, UPS, Université de Toulouse
- Pascal Perrier, GIPSA-lab, Université de Grenoble
- Myriam Piccaluga, Laboratoire de Phonétique, Université de Mons
- Claire Pillot-Loiseau, LPP, Université Sorbonne Nouvelle
- Julien Pinquier, IRIT, UPS, Université de Toulouse

- Halima Sahraoui, LNPL, UT2J, Université de Toulouse
- Jean Schoentgen, BEAMS, Université Libre de Bruxelles
- Rudolph Sock, LILPA, Université de Strasbourg
- Béatrice Vaxelaire-Sock, LILPA, Université de Strasbourg
- Virginie Woisard, Hôpitaux de Toulouse, LNPL, UT2J, Université de Toulouse
- Naomi Yamaguchi, ILPGA & LPP, Université Sorbonne Nouvelle

# Comité d'organisation

- Corine Astésano, LNPL, UT2J, Université de Toulouse
- Mathieu Balaguer, IRT, UPS, CHU/Université de Toulouse
- Jérôme Farinas, IRT, UPS, Université de Toulouse
- Corinne Fredouille, LIA, Avignon Université
- Lila Gravelier, IRT, UPS, Université de Toulouse
- Fanny Guitard-Ivent, LNPL, UT2J, Université de Toulouse
- Anna Marczyk-Buklaha, LNPL, UT2J, Université de Toulouse
- Julie Maclair, IRT, UPS, Université de Toulouse
- Tuan Mahn Nguyen, LIA, Avignon Université
- Julien Pinquier, IRT, UPS, Université de Toulouse
- Halima Sahraoui, LNPL, UT2J, Université de Toulouse
- Virginie Woisard, Hôpitaux de Toulouse, LNPL UT2J, Université de Toulouse

# Communications invitées



# Phénomènes compensatoires : utilité pour l'évaluation et le traitement de la parole déviante

Lorraine BAQUÉ

En tant qu'êtres éminemment adaptatifs, les humains font face à leurs difficultés de communication en développant – spontanément ou de manière induite – des stratégies compensatoires plus ou moins efficaces. De tels comportements sont omniprésents dans la communication humaine, en production comme en réception, mais leur utilité dans le cadre de la rééducation des troubles du langage ou de la parole a fait l'objet de nombreuses controverses.

Dans cette présentation, nous rapporterons le résultat de recherches sur des stratégies compensatoires au plan phonique développées par des patients souffrant de troubles de la parole et du langage, avant d'en évaluer l'efficacité pour l'intelligibilité segmentale et prosodique. Nous concluons par une réflexion sur le rôle que peut avoir l'analyse de ces stratégies à la fois pour l'évaluation des déficits et pour la prise en charge orthophonique.

---

## Biographie

Professeur de phonétique et de linguistique française à l'Universitat Autònoma de Barcelona (Espagne) et orthophoniste, Lorraine Baqué s'est spécialisée dans l'analyse de la parole déviante au niveau segmental et prosodique. Sa recherche porte plus particulièrement sur des pathologies du langage et de la parole d'étiologie neurologique, telles l'aphasie, l'apraxie de la parole et l'aphasie progressive primaire, ainsi que sur l'acquisition de langues secondes ou étrangères.



# Co-designing smart technologies for self-expression: case studies with people with intellectual disability

Laurianne SITBON

Bridging the gap between advances in artificial intelligence and assistive technologies, I will discuss why and how we can co-design pictorial visual communication applications with people with intellectual disability or autism and community members. Through this discussion, I will illustrate the approaches we have developed based on individuals and their competencies, such as self-expression by design and competency-based design. I will then showcase some of the work we have conducted to co-design technologies that can encourage self-expression and amplify the voices of people with intellectual disability, including when they participate in research. I will particularly emphasise the potential for online materials to offer virtually unlimited resources to support communication and connection with others. Finally, I will offer some reflections on how voice-based interactive systems can adapt to emphasise the voices of people with intellectual disability when they are supported to use them, and how multi-modality may account for people's diverse ways of expressing themselves verbally.

---

## Biographie

Laurianne is a Future Fellow of the Australian Research Council (ARC), and Associate Professor in the School of Computer Science at the Queensland University of Technology (QUT) in Brisbane, Australia. She is co-lead of the Human Interaction program of the QUT Centre for Robotics, chief investigator in the QUT Centre for Justice, and Academic Lead Diversity and Inclusion in the School of Computer Science.

Her research expertise now spans human computer interactions (with a focus on co-design and cognitive accessibility), natural language processing (with a focus on semantics), and information retrieval. Her current projects include the requirements of multi-modal conversational systems that can best support the needs of people with intellectual disability (<https://research.qut.edu.au/dplab/research/a-pictorial-communication-framework-for-inclusion/>), and more generally how people, such as experts, can “converse” with systems, which analyse data and make predictions, to assist them in making decisions (<https://research.qut.edu.au/dplab/research/human-machine-teaming>). She obtained her PhD in Computer Science from the University of Avignon in France, in 2008.



# **L'analyse acoustique de la voix en contexte clinique : recommandations actuelles et enjeux cliniques**

Vincent MARTEL-SAUVAGEAU

L'analyse acoustique de la voix est une approche largement répandue dans le contexte de la recherche. Bien que variable selon les milieux, son utilisation en clinique est maintenant en croissance, permettant aux cliniciens de compléter leurs mesures perceptuelles classiques, et ce tant pour l'évaluation que dans une approche de rééducation.

Néanmoins, la réalité clinique impose des défis pratiques indéniables pour l'utilisation des mesures acoustiques auprès des clients présentant des troubles de la voix. Cette présentation abordera les différents défis liés à la capture du signal vocal en contexte clinique et présentera certaines recommandations d'application, notamment sur la prise de mesure, la calibration du signal et les tâches d'évaluations. Les enjeux et recommandations concernant l'analyse du signal seront également abordés, en termes de variables acoustiques et de leur interprétation, ainsi que concernant différents logiciels d'analyse.

---

## **Biographie**

Vincent Martel-Sauvageau est orthophoniste et professeur agrégé au programme d'orthophonie de l'université Laval (Québec, Canada). Ses principaux travaux de recherche portent sur les propriétés acoustiques fines de la parole et de la voix saine et pathologique, de même que sur les marqueurs de l'intelligibilité de la parole. Il s'intéresse également à l'amélioration des pratiques orthophoniques d'évaluation et d'intervention des troubles de la parole en contexte franco-québécois, de même qu'aux aspects méthodologiques associés à l'analyse du signal acoustique. Il est affilié au Centre interdisciplinaire de recherche en réadaptation et intégration sociale (CIRRIIS), à Québec (Canada).



# Communications orales



# Voix, parole et surdit 



# Échelles perceptives pour les voix parlées normophoniques, pathologiques et chantées : revue de littérature et réflexions méthodologiques

Amelia PETTIROSSI<sup>1,3</sup>  
Laura WESTERMEYER<sup>1</sup>  
Amira ROUCHEZ SLIMANI<sup>1</sup>  
Nicolas AUDIBERT<sup>1</sup>  
Lise CREVIER-BUCHMAN<sup>1,4</sup>  
Véronique DELVAUX<sup>2</sup>  
Sophie FAGNIART<sup>2</sup>  
Bernard HARMEGNIES<sup>2</sup>  
Kathy HUET<sup>2</sup>  
Myriam PICCALUGA<sup>2</sup>  
Reina REMMAN<sup>1</sup>  
Virginie ROLAND<sup>2</sup>  
Lola TERNY<sup>2</sup>  
Clémence VERHAEGEN<sup>2</sup>  
Claire PILLOT-LOISEAU<sup>1</sup>

<sup>1</sup>Laboratoire de Phonétique et phonologie UMR 7018, CNRS/Université Sorbonne Nouvelle, Paris, France

<sup>2</sup>Service de Métrologie et Sciences du Langage, Institut de Recherche en Sciences et Technologies du Langage & Fond National de La Recherche Scientifique, Université de Mons, Belgique

<sup>3</sup>Imagerie & Cerveau, UMR1253 INSERM/Université de Tours, 37000 Tours, France

<sup>4</sup>Hôpital Foch : Service de Laryngologie Phoniatrie, Université Paris Saclay, Suresnes

Cette revue de littérature préliminaire de 26 articles s'est consacrée à recenser les échelles perceptives utilisées par les (socio)phonéticiens, pathologistes ou pédagogues du chant pour l'hétéroévaluation des voix parlées normophoniques (Fagel et al., 1983 ; Gelfer 1988), dysphoniques (Pinho & Pontes, 2002 ; Yamasaki et al., 2017) et chantées (Oates et al., 2006 ; Henrich et al., 2008). Les limites et atouts de différents instruments perceptifs d'hétéroévaluation vocale seront mis en évidence afin de préciser les propriétés métrologiques et les précautions méthodologiques qui leurs sont associées.

Nous relevons de nos lectures plusieurs types d'instruments perceptifs, dont la majorité investiguent la qualité vocale globale sans en considérer les variations locales. Il existe des échelles visuelles analogiques à l'image du *CAPE-V* (Kempster et al., 2009), des échelles ordinales avec peu de paramètres à évaluer comme le *GRBAS* (Hirano, 1981), des échelles ordinales avec de nombreux

paramètres tel que le *VPAS* (Laver et al., 1981) mais aussi diverses échelles ordinales basées sur le principe d'échelles sémantiques différentielles (Osgood, 1952). La plupart des échelles ordinales ont des intervalles uniquement numérotés.

Nous notons que les accords inter-juges sont rarement élevés, surtout lorsqu'on fait appel à des paramètres plus subjectifs comme *unpleasant-pleasant* (Fagel et al., 1983) qui est un élément d'appréciation esthétique et non un paramètre vocal. En revanche, les échelles ordinales obtiennent un meilleur accord inter-juges que celles visuelles analogiques.

De plus, la plupart sont créées ou utilisées en langue anglaise ce qui peut poser des problèmes de traduction de certains items (*hoarse, husky* ou *rough*, sont tous traduits par *rauque*).

Nous constatons également que les consignes et paramètres perceptifs évalués sont grandement dépendants du type de production (voix chantées, pathologiques ou normales), du corpus (voyelle tenue ou parole continue), du profil des auditeurs (entraînés ou non, experts ou naïfs) et du but de l'investigation (recherche de norme pour les productions pathologiques ou satisfaction esthétique pour les productions chantées).

Plusieurs recommandations méthodologiques découlent de nos observations : soigner les présentations visuelles (ne pas mettre tous les termes négatifs à gauche en cas d'échelle sémantique différentielle, ne pas mélanger les items normaux et pathologiques) ; bien expliciter les paramètres et le type de corpus présenté, prendre en considération le profil des auditeurs et adopter un compromis entre un nombre suffisant et un nombre raisonnable de paramètres pour que leur traitement perceptif soit fiable et valide. Enfin, il faudrait envisager des exemples sonores de référence comme ancrage.

### Références bibliographiques

- Fagel, W. P. F., Van Herpt, L. W. A., & Boves, L. (1983). Analysis of the perceptual qualities of Dutch speakers' voice and pronunciation, *Speech Communication*, 2(4), 315-326.
- Gelfer, M. P. (1988). Perceptual attributes of voice: Development and use of rating scales, *Journal of Voice*, 2(4), 320-326.
- Henrich, N., Bezar, P., Expert, R., Garnier, M., Guerin, C., Pillot-Loiseau, C., Quattrocchi, S., Roubeau, B., & Terk B. (2008). Towards a common terminology to describe voice quality in western lyrical singing: Contribution of a multidisciplinary research group, *Journal of interdisciplinary music studies*, 2(1&2), 71-93.
- Hirano, M. (1981). Clinical examination of voice, *Disorders of human communication*, 5, 1-99.
- Kempster, G. B., Gerratt, B. R., Verdolini Abbott, K., Barkmeier-Kraemer, J., & Hillman, R. E. (2009). Consensus auditory-perceptual evaluation of voice: development of a standardized clinical protocol, *American Journal of Speech Language Pathology*, 18, 124-132.
- Laver, J., Wirz, S., Mackenzie, J., et al. (1981). A perceptual protocol for the analysis of vocal profiles, *Edinburgh University Department of Linguistics Work in Progress*, 14, 139-155.
- Oates, J., Bain, B., Davis, P. B., Chapman, J. A., & Kenny, D. T. (2006). Development of an Auditory-Perceptual Rating Instrument for the Operatic Singing Voice. *Journal of Voice*, 20(1), 71-81. <https://doi.org/10.1016/j.jvoice.2005.01.006>
- Osgood, C. E. (1952). The nature and measurement of meaning, *Psychological bulletin*, 49(3), 197.

- Pinho, S. M. R., & Pontes, P., (2002). Escala de avaliação perceptiva da fonte glótica: RASAT, *Vox Brasiliis*, 3(1), 11-3.
- Yamasaki, R., Madazio, G., Leão, S. C., Padovani, M., Azevedo, R. C., & Behlau, M. (2017). Auditory-perceptual Evaluation of Normal and Dysphonic Voices Using the Voice Deviation Scale. *Journal of Voice*, 31(1), 67-71. <https://doi.org/10.1016/j.jvoice.2016.01.004>

### **Remerciements**

Ce travail a été soutenu par le Partenariat Hubert Curien (PHC) TOURNESOL n°46241WH et par le Laboratoire d'Excellence (LabEx) *Empirical Foundations of Linguistics* (EFL) n°ANR-10-LABX-0083. Il contribue à l'IdEx Université de Paris (ANR-18-IDEX-0001).



# Contribution de l'information segmentale dans l'accès lexical chez des enfants avec implants cochléaires et enfants avec audition typique

Bénédicte GRANDON<sup>1</sup>  
Esther RUIGENDIJK<sup>1</sup>

<sup>1</sup> Institut d'Etudes Néerlandaises &  
Cluster d'Excellence Hearing4all,  
Université d'Oldenburg, Allemagne

Dès la fin de leur première année, les enfants sont capables de catégoriser les sons de leur langue maternelle à partir de leurs caractéristiques acoustiques (e.g., Jusczyk, 1992 ; Pierrehumbert, 2001) et ils sont capables de construire des représentations lexicales en utilisant l'information disponible aux niveaux segmental, syllabique ou prosodique (Nazzi, 2006). Ce processus peut cependant être perturbé dans le cas d'une surdité. L'implant cochléaire (IC) permet à ses utilisateurs d'accéder à une information auditive plus riche, même si le signal transmis par l'implant reste incomplet (Drennan & Rubinstein, 2008), avec des conséquences sur le traitement de certains sons et des difficultés à plus haut niveau de traitement linguistique (e.g., lexique, morphologie, syntaxe...). Ainsi, l'accès lexical chez les enfants avec IC reste plus compliqué et tardif que chez les enfants avec audition typique (Schwartz et al., 2013), et avec une plus grande variabilité interlocuteurs (e.g., adultes : Nagels et al., 2020). Cependant, le rôle des caractéristiques des sons qui constituent les mots dans le processus d'accès lexical en réception reste peu étudié. Notre étude s'intéresse à l'utilisation des caractéristiques segmentales des mots dans l'accès lexical en réception par des enfants avec IC. Nous faisons l'hypothèse qu'en raison de leurs caractéristiques intrinsèques, tous les sons ne sont pas aussi saillants les uns que les autres, ce qui a pour conséquence des difficultés à accéder au mot, en particulier lorsque toute l'information acoustique n'est pas disponible, dans le cas d'un traitement par l'IC.

Utilisant un paradigme comportemental et d'eye-tracking, notre étude proposait à deux groupes d'enfants germanophones de 5 à 11 ans (28 enfants avec audition typique, 15 enfants avec IC dont 3 unilatéraux, âge à l'implantation : 0:6-1;10) d'identifier un mot présenté oralement parmi quatre images (mot-cible, compétiteur phonologique, et deux distracteurs sémantiques). Le mot-cible et le compétiteur phonologique se différenciaient d'un à trois traits d'une consonne ou d'une voyelle (voir *Tableau 1*). Nous avons mesuré la justesse et le temps de réaction des réponses, ainsi que l'évolution temporelle du pourcentage de fixation vers l'image cible.

Les résultats indiquent une meilleure identification des mots-cibles et une fixation plus rapide du mot-cible chez les enfants avec audition typique, et une différence de traitement selon le nombre mais pas selon le type de différences entre cible et compétiteur, ce qui nous laisse envisager un traitement qualitatif similaire des indices segmentaux par les deux groupes d'enfants, mais un processus de traitement plus lent des enfants avec IC pour accéder au mot.

Tableau 1. Types et nombres de traits contrastifs considérés

Segment cible	Nombre de traits contrastifs	Traits	Exemple (cible-compétiteur phonologique)
Consonne	1	mode d'articulation	/pas/-/fas/, <i>passerport-tonneau</i>
		lieu d'articulation (fricatives)	/vant/-/zant/, <i>mur-sable</i>
		lieu d'articulation (plosives)	/bɛilə/-/gɛilə/, <i>lunettes-grillon</i>
		voisement	/kabəl/-/gabəl/, <i>cable-fourchette</i>
	2	mode + voisement	/bʊtə/-/fʊtə/, <i>beurre-nourriture</i>
		lieu + mode	/tɪʃ/-/fɪʃ/, <i>table-poisson</i>
		lieu + voisement (plosives)	/tu:χ/-/bu:χ/, <i>serviette-livre</i>
3	lieu + mode + voisement	/be:ʁə/-/ʃe:ʁə/, <i>baie-ciseaux</i>	
Voyelle	1	antériorité	/ba:t/-/bɛt/, <i>baignoire-lit</i>
		hauteur	/mʊnt/-/mo:nt/, <i>bouche-lune</i>
		arrondissement	/kɪstə/-/kʏstə/, <i>caisse-côte</i>
	2	antériorité + hauteur	/rybə/-/rɔbə/, <i>navet-phoque</i>
		antériorité + arrondissement	/kɔfə/-/kɛ:fə/, <i>scarabée-valise</i>

### Références bibliographiques

- Drennan, W. R., & Rubinstein, J. T. (2008). Music perception in cochlear implant users and its relationship with psychophysical capabilities. *Journal of rehabilitation research and development*, 45(5), 779.
- Jusczyk, P. W. (1992). Developing phonological categories from the speech signal. *Phonological development: Models, research, implications*, 17-64.
- Nagels, L., Bastiaanse, R., Başkent, D., & Wagner, A. (2020). Individual differences in lexical access among cochlear implant users. *Journal of Speech, Language, and Hearing Research*, 63(1), 286-304.
- Nazzi, T. (2006). Spécificité phonétique : de la perception précoce à l'acquisition des premiers mots. *Recherches linguistiques de Vincennes*, (35), 31-52.
- Pierrehumbert, J. B. (2003). Phonetic diversity, statistical learning, and acquisition of phonology. *Language and speech*, 46(2-3), 115-154.
- Schwartz, R. G., Steinman, S., Ying, E., Mystal, E.Y., Houston, D. M. (2013). Language processing in children with cochlear implants: a preliminary report on lexical access for production and comprehension. *Clinical linguistics & phonetics*, 27(4), 264-277.

# Amélioration de la production de consonnes chez les enfants porteurs d'implants cochléaires lorsqu'ils sont exposés à la Langue française Parlée Complétée : corrélats acoustiques et articulatoires

Laura MACHART<sup>1,2</sup>  
Anne VILAIN<sup>1</sup>  
Hélène L'ÉVENBRUCK<sup>2</sup>  
Mark TIEDE<sup>3</sup>  
Lucie MÉNARD<sup>4</sup>

<sup>1</sup>GIPSA-lab,  
UMR 5216 CNRS,  
Université Grenoble Alpes - Grenoble INP

<sup>2</sup>Laboratoire de Psychologie et NeuroCognition,  
UMR 5105 CNRS,  
Université Grenoble Alpes

<sup>3</sup>Haskins Laboratories,  
New Haven, Connecticut, USA

<sup>4</sup>Laboratoire de Phonétique,  
Université du Québec à Montréal

La Langue française Parlée Complétée (LfPC) est un code visuo-manuel qui soutient la communication en langue vocale chez l'enfant sourd. Les gestes manuels sont utilisés comme des informations phonologiques supplémentaires qui complètent les informations acoustiques et labiales de la parole. De nombreuses études ont démontré que l'exposition régulière à la LfPC améliore la perception de parole et le développement des compétences phonologiques de l'enfant sourd (Leybaert et al., 2016 ; Leybaert & LaSasso, 2010). Il a également été suggéré que cette exposition améliore la production de parole chez l'enfant porteur d'implants cochléaires (Hage & Leybaert, 2005; Machart et al., 2020, soumis).

L'implant cochléaire (IC) favorise l'intelligibilité de la parole des enfants sourds (Grandon et al., 2020) mais les informations auditives transmises demeurent dégradées (Colin et al., 2017), ce qui peut engendrer des troubles du langage persistants (Geers et al., 2016). Des difficultés de production de parole chez les enfants francophones porteurs d'IC ont été mises en évidence par des mesures acoustiques (Grandon, 2016) et articulatoires (Turgeon et al., 2017). L'objectif de cette étude est d'examiner l'influence du niveau de décodage de la LfPC, c'est-à-dire la capacité à interpréter ce code visuo-manuel, sur la précision articulatoire consonantique d'enfants porteurs d'IC à partir de données ultrasons et acoustiques.

Les productions de 18 enfants âgés de 52 à 137 mois ont été étudiées : huit enfants porteurs d'IC (mean = 102.55 mois, sd = 19.35) avec différents niveaux de décodage de la LfPC sont comparés à dix enfants au développement typique (mean = 96.25 mois, sd = 25.68). Des paramètres acoustiques ont été mesurés pour les plosives /t/ et /k/ et fricatives /s/ et /ʃ/

(formants et premier moment spectral). Les contours de langue ont été traités avec le programme SLURP (Laporte & Ménard, 2018) et ont été évalués en fonction de l'index de courbure MCI (Dawson et al., 2016), de la position horizontale de la langue (Ménard et al., 2013) et du lieu de constriction (Ohkubo & Scobbie, 2019).

Les résultats préliminaires de l'analyse statistique avec un modèle linéaire mixte avec effet aléatoire participant révèlent qu'un niveau élevé de décodage de la LfPC est bénéfique à la production de consonnes chez les enfants porteurs d'IC, quel que soit leur seuil auditif post-implantation (amélioration du lieu d'articulation des plosives et fricatives). Ils confirment également la nécessité de comparer mesures acoustiques et articulatoires pour mieux caractériser la production de parole chez les enfants.

### Références bibliographiques

- Colin, S., Ecalte, J., Truy, E., Lina-Granade, G., & Magnan, A. (2017). Effect of age at cochlear implantation and at exposure to Cued Speech on literacy skills in deaf children. *Research in Developmental Disabilities, 71*, 61–69.
- Dawson, K. M., Tiede, M. K., & Whalen, D. H. (2016). Methods for quantifying tongue shape and complexity using ultrasound imaging. *Clinical Linguistics & Phonetics, 30*(3–5), 328–344.
- Geers, A. E., Nicholas, J., Tobey, E., & Davidson, L. (2016). Persistent Language Delay Versus Late Language Emergence in Children With Early Cochlear Implantation. *Journal of Speech, Language, and Hearing Research, 59*(1), 155–170.
- Grandon, B. (2016). *Développement typique et atypique de la production de parole: Caractéristiques segmentales et intelligibilité de la parole d'enfants porteurs d'un implant cochléaire et d'enfants normo-entendants de 5 à 11 ans*. Université Grenoble Alpes.
- Grandon, B., Martinez, M.-J., Samson, A., & Vilain, A. (2020). Long-term effects of cochlear implantation on the intelligibility of speech in French-speaking children. *Journal of Child Language, 47*(4), 881–892.
- Hage, C., & Leybaert, J. (2005). The Effect of Cued Speech on the Development of Spoken Language. In P. E. Spencer & M. Marschark (Eds.), *Advances in the Spoken Language Development of Deaf and Hard-of-Hearing Children* (pp. 193–211). Oxford University Press.
- Laporte, C., & Ménard, L. (2018). Multi-hypothesis tracking of the tongue surface in ultrasound video recordings of normal and impaired speech. *Medical Image Anal.*
- Leybaert, J., Bayard, C., Colin, C., & LaSasso, C. (2016). Cued Speech and Cochlear Implants: A Powerful Combination For Natural Spoken Language Acquisition and the Development of Reading. In M. Marschark & P. E. Spencer (Eds.), *The Oxford Handbook of Deaf Studies in Language*. Oxford University Press.
- Leybaert, J., & LaSasso, C. J. (2010). Cued Speech for Enhancing Speech Perception and First Language Development of Children With Cochlear Implants. *Trends in Amplification, 14*(2), 96–112.

- Machart, L., Vilain, A., Løevenbruck, H., Gerber, S., & EULALIES Consortium. (soumis). Positive impact of French Cued Speech exposure on speech production in children fitted with Cochlear Implants. *PlosOne*.
- Machart, L., Vilain, A., Løevenbruck, H., Meloni, G., & Puissant, C. (2020). Production de parole chez l'enfant porteur d'implant cochléaire: Apport de la Langue française Parlée Complétée. *JEP-TALN-RECITAL 2020-6e Conférence Conjointe 33e Journées d'Études Sur La Parole, 27e Traitement Automatique Des Langues Naturelles, 22e Rencontre Des Étudiants Chercheurs En Informatique Pour Le Traitement Automatique Des Langues*, 388–396.
- Ménard, L., Toupin, C., Baum, S. R., Drouin, S., Aubin, J., & Tiede, M. (2013). Acoustic and articulatory analysis of French vowels produced by congenitally blind adults and sighted adults. *The Journal of the Acoustical Society of America*, 134(4), 2975–2987.
- Ohkubo, M., & Scobbie, J. M. (2019). Tongue Shape Dynamics in Swallowing Using Sagittal Ultrasound. *Dysphagia*, 34(1), 112–118.
- Turgeon, C., Trudeau-Fisette, P., Pitzpatrick, E., & Ménard, L. (2017). Vowel intelligibility in children with cochlear implants: An acoustic and articulatory study. *International Journal of Pediatric Otorhinolaryngology*, 101, 87–96.



# Nasalité vocalique, compétences phonologiques et grammaticales : des liens spécifiques chez les enfants porteurs d'implants cochléaires ?

Sophie FAGNIART<sup>1</sup>  
Véronique DELVAUX<sup>1,2</sup>  
Brigitte CHARLIER<sup>3,4</sup>  
Bernard HARMEGNIES<sup>1</sup>  
Anne HUBERLANT<sup>3</sup>  
Myriam PICCALUGA<sup>1</sup>  
Kathy HUET<sup>1</sup>

<sup>1</sup> Service de Métrologie et Sciences du Langage, Institut de Recherche en Sciences et Technologies du Langage, UMONS, Belgique

<sup>2</sup> Fonds National de la Recherche Scientifique, Belgique

<sup>3</sup> Centre Comprendre et Parler,  
Bruxelles

<sup>4</sup> Université Libre de Bruxelles,  
Bruxelles

De nombreuses études ont rapporté des difficultés morphosyntaxiques chez les enfants porteurs d'implant(s) cochléaire(s) (Bourdin et al., 2016 ; Duchesne, 2015 ; Lenormand, 2004). Les limitations de perception auditive associées à des représentations phonologiques imprécises et pouvant conduire à des difficultés de traitement de morphèmes grammaticaux, peu saillants dans la langue parlée, sont notamment évoquées comme pistes explicatives à ces difficultés (Guo & Spencer, 2017 ; Hansson et al., 2017). Aucune étude n'a toutefois investigué cette piste en langue française, où de nombreuses marques grammaticales sont portées par des traits acoustiques fins, notamment par l'opposition entre voyelles nasales et orales problématique au sein de la population implantée (Fagniard et al., 2021 ; Borel, 2019 ; Bouton et al., 2012).

Dans ce contexte, notre étude présente les analyses préliminaires menées sur 10 enfants implantés (groupe IC) âgés de 4;7 à 7;10 ans et 15 enfants normo-entendants (groupe NE) âgés de 2;11 à 6;2 ans. Les enfants ont réalisé différentes tâches langagières, dont une tâche de dénomination d'images et deux productions de récits. Pour la tâche de dénomination d'images, nous avons calculé un pourcentage de phonèmes corrects (utilisé pour apprécier les compétences phonologiques) et conduit des analyses acoustiques sur les productions. Des distances entre paires de voyelles nasales et orales ont été calculées sur base de deux indices acoustiques : les valeurs de A1-P0 (Chen, 1997), reflétant la mise en résonance nasale, et les distances euclidiennes sur les plans F1/F2/F3, reflétant principalement la configuration articulaire oro-pharyngée. Ces choix méthodologiques sont basés sur des résultats antérieurs ayant indiqué une utilisation préférentielle des indices de configuration articulaire dans le marquage de la nasalité vocalique chez les enfants IC (Fagniard et al., 2021). Les récits ont été retranscrits et un étiquetage syntaxique automatique des mots a été réalisé à l'aide du logiciel CLAN. Certaines classes de mots ont été regroupées afin d'obtenir un indice de morphologie grammaticale (Lenormand, 2004).

D'une part, les résultats montrent des compétences linguistiques plus faibles dans le groupe IC, conformément à la littérature.

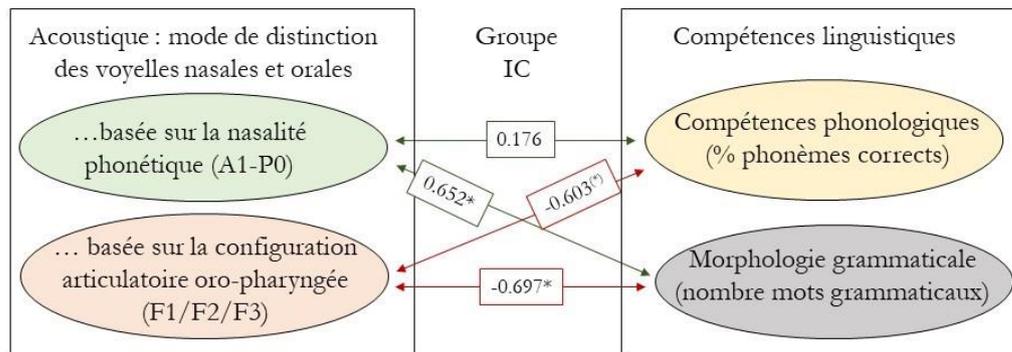


Figure 1. Corrélations entre nos deux types d'indices acoustiques (A1-P0 et distances euclidiennes F1/F2/F3), les compétences phonologiques (% de phonèmes corrects) et grammaticales (indice de morphologie grammaticale, Lenormand, 2004), au sein du groupe IC. Les corrélations ayant un niveau de significativité  $<.05$  sont marquées par \* et  $<.08$  marquées par (\*).

D'autre part, nous observons des liens différenciés (corrélations positives vs. négatives) entre nos deux types d'indices acoustiques et les compétences linguistiques associées au sein du groupe IC (Figure 1), ce qui n'est pas le cas dans le groupe NE (Figure 2).

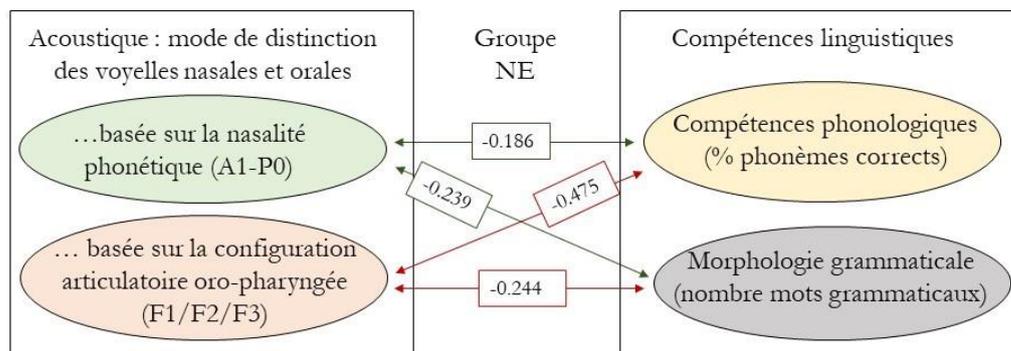


Figure 2. Corrélations entre nos deux types d'indices acoustiques (A1-P0 et distances euclidiennes F1/F2/F3), les compétences phonologiques (% de phonèmes corrects) et grammaticales (indice de morphologie grammaticale, Lenormand, 2004) au sein du groupe NE. Aucune corrélation n'atteint un niveau de significativité  $<.08$ .

Ces résultats seront discutés à la lumière des théories phonologiques des difficultés linguistiques (« mapping theory », Chiat, 2001).

### Références bibliographiques

Borel, S., Serniclaes, W., Sterkers, O., & Vaissière, J. (2019). Identification of nasal consonants and nasal vowels by French adults cochlear implant listeners. *Audiology Direct*, 3(1), 1-7.

- Bourdin, B., Ibernon, L., Le Driant, B., Levrez, C., & Vandromme, L. (2016). Troubles morphosyntaxiques chez l'enfant sourd et chez l'enfant dysphasique: similarités et spécificités. *Revue de neuropsychologie*, 8(3), 161-172.
- Bouton, S., Serniclaes, W., Bertoncini, J., & Colé, P. (2012). Perception of speech features by French-speaking children with cochlear implants. *Journal of Speech, Language, and Hearing Research*, 55, 139–153
- Chen, M. Y. (1997). Acoustic correlates of English and French nasalized vowels. *The Journal of the Acoustical Society of America*, 102(4), 2360-2370.
- Chiat, S. (2001). Mapping theories of developmental language impairment: Premises, predictions and evidence. *Language and cognitive processes*, 16(2-3), 113-142.
- Duchesne, L. (2010). *Développement linguistique d'enfants porteurs d'un implant cochléaire: le vocabulaire et la grammaire chez les enfants ayant reçu un implant en bas âge*. Éditions universitaires européennes.
- Fagniat, S., Charlier, B., Delvaux, V., Huberlant, A., Huet, K., Piccaluga, M., Watterman, I., & Harmegnies, B. (2020). *Perception et production du trait de nasalité vocalique chez l'enfant porteur d'implants cochléaires*. In 6<sup>e</sup> Conférence Conjointe Journées d'Études sur la Parole (JEP, 33<sup>e</sup> Édition). Nancy, 217-225.
- Guo, L. Y., & Spencer, L. J. (2017). Development of grammatical accuracy in English-speaking children with cochlear implants: A longitudinal study. *Journal of Speech, Language, and Hearing Research*, 60(4), 1062-1075.
- Hansson, K., Ibertsson, T., Asker-Árnason, L., & Sahlén, B. (2017). Phonological processing, grammar and sentence comprehension in older and younger generations of Swedish children with cochlear implants. *Autism & Developmental Language Impairments*, 2, 1-14.
- Le Normand, M. T. (2004). Évaluation du lexique de production chez des enfants sourds profonds munis d'un implant cochléaire sur un suivi de trois ans. *Rééducation orthophonique*, 217, 125-140.



# Analyse Factorielle des Enveloppes de Modulations d'Amplitude des signaux de la parole dans la perspective de l'hypothèse du codage efficace et de l'application aux implants cochléaires

Agnieszka DUNIEC<sup>1</sup>  
Olivier CROUZET<sup>1</sup>  
Elisabeth DELAIS-ROUSSARIE<sup>1</sup>

<sup>1</sup>Laboratoire de Linguistique de Nantes,  
UMR 6310,  
Université de Nantes - CNRS

L'hypothèse du « codage efficace » (Smith & Lewicki, 2006) postule que les signaux acoustiques de communication sont caractérisés par des propriétés statistiques régulières et que les systèmes perceptifs sont optimalement adaptés à ces signaux de manière à transmettre un maximum d'information en recourant à une consommation minimale de ressources. Ainsi, certains travaux se sont centrés sur (1) le nombre optimal de canaux spectraux pour représenter des langues orales mais aussi (2) la localisation des frontières entre ces canaux. Les changements de localisation des frontières spectrales en parole vocodée aurait des effets sur les taux de reconnaissance des mots (Ming & Holt, 2009). 4 canaux suffiraient à représenter optimalement des signaux de parole pour 8 langues différentes (Ueda & Nakajima, 2017). Il faudrait 6 à 7 canaux pour représenter de manière appropriée la parole vocodée dans des simulations d'implants cochléaires (Grange & Culling, 2018).

Nous avons mis en place une série d'analyses afin de décrire des propriétés statistiques de signaux naturels de parole et les comparer aux travaux antérieurs.

Les analyses acoustiques et statistiques ont été réalisées dans l'environnement Matlab et reproduisent les travaux précédents (Ueda & Nakajima, 2017; Grange & Culling, 2018). La base de données est un corpus de parole en langue anglaise, Clarity Speech (Graetzer et al., 2022). Les enveloppes de modulation temporelle des signaux ont été extraites à partir d'un banc de filtres gammatone (environ 120 canaux). La matrice de données résultante a été transférée vers un outil statistique (ACP) afin de procéder à une Analyse Factorielle.

Dans l'analyse réalisée par Ueda & Nakajima (2017), les frontières entre canaux décrits comme optimaux sont placées aux croisements des courbes de saturation, délimitant ainsi des zones de fréquences. Dans la Fig. 1, nous avons représenté les frontières proposées par les auteurs à travers des traits verticaux de couleur orange. L'analyse proposée par les auteurs identifie 3 frontières qui délimitent quatre zones de fréquence principales. Si l'on observe les données de courbes de saturation pour 4 composantes principales ( $n_{PC}=4$ ), on voit clairement les 4 courbes délimitant des zones de fréquences différentes, conformément aux résultats obtenus par Ueda & Nakajima (2017). Toutefois, les frontières optimales proposées par les auteurs ne correspondent pas, c'est-à-dire, ne sont pas placées aux croisements des courbes, notamment en ce qui concerne la frontière entre les zones de fréquences moyennement hautes et hautes. Le graphe des valeurs propres (Fig. 2) représente le cumul de variance expliquée pour les facteurs retenus. Si l'on se restreint aux 22 premiers facteurs, on atteint un cumul de 88 % de variance expliquée, 20 points de plus que Grange & Culling, 2018. Pour 4-7 facteurs, ces valeurs sont de 30% à 50%, 10 à 15 points

de plus. Nous finalisons l'automatisation du calcul du nombre optimal des Composantes Principales ainsi que des frontières de fréquence. Des études perceptives permettront d'évaluer la validité de ces frontières pour la perception de l'intonation et de la  $f_0$ .

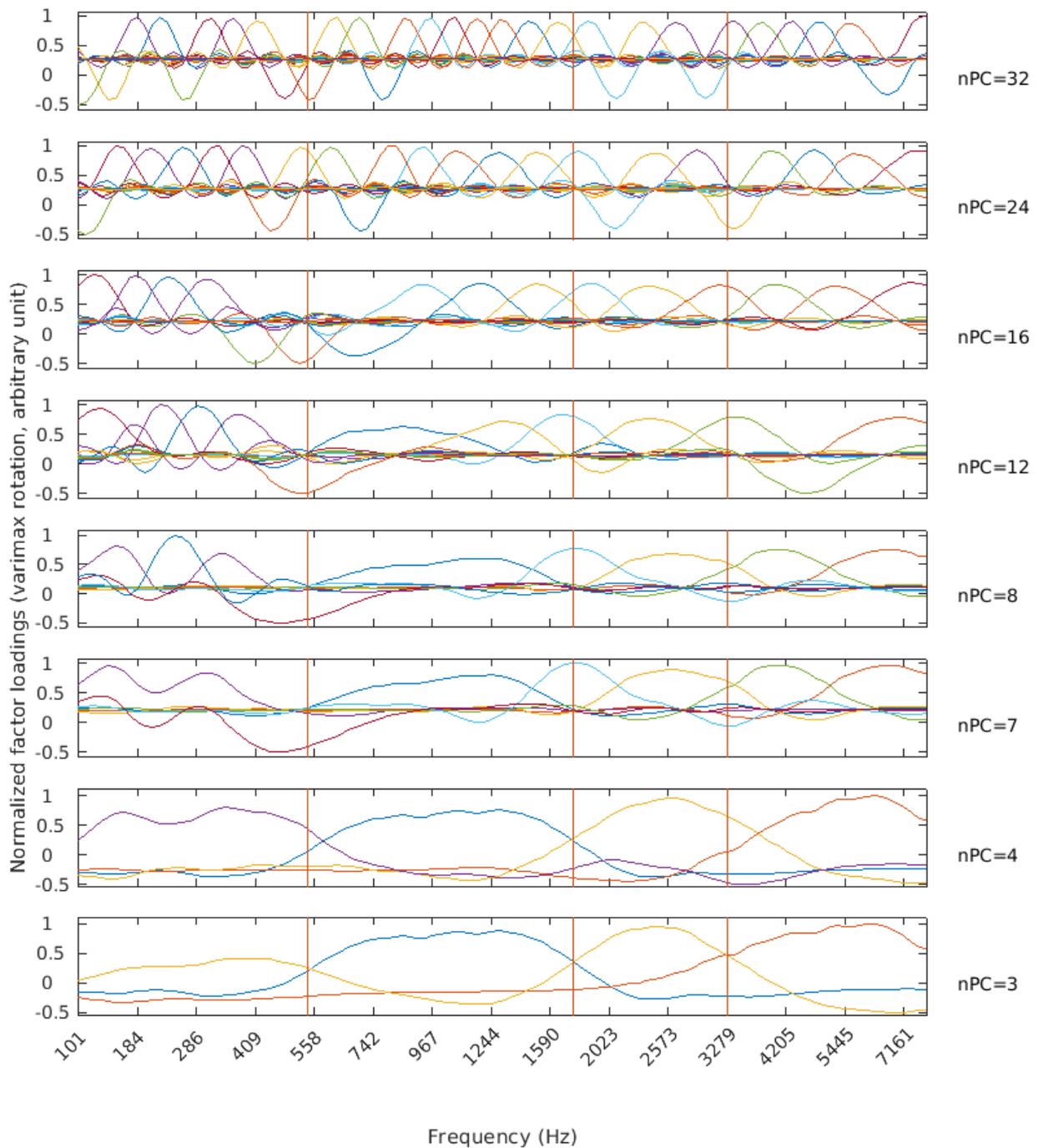


Figure 1. Courbes de saturations issues de l'échantillon du signal de parole d'une durée de 1 heure (extraits aléatoires) ; fréquence max 8kHz.

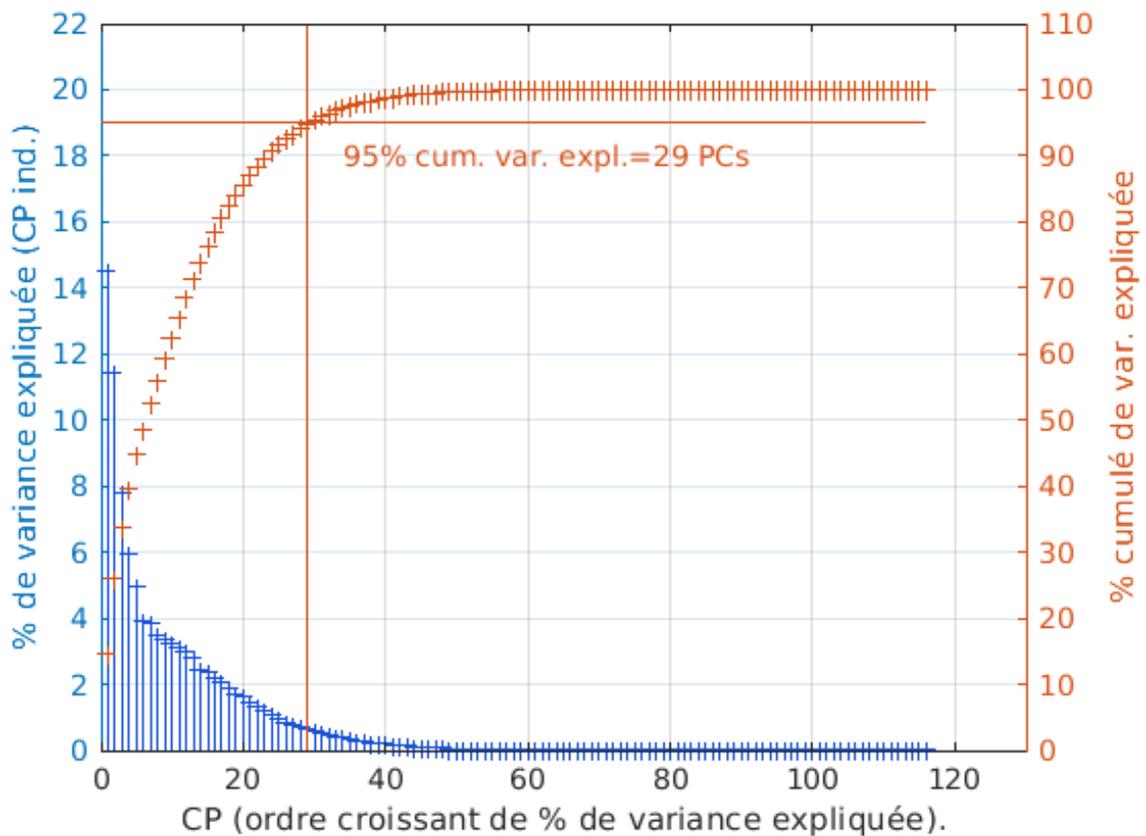


Figure 2. Graphe des valeurs propres ou screen plot (en bleu : valeurs individuelles de variance expliquée, en rouge : pourcentage cumulé de variance expliquée) pour 1 heure de signal de parole ; fréquence max 8khz.

### Références bibliographiques

- Graetzer, S., Akeroyd, M. A., Barker, J., Cox, T. J., Culling, J. F., Naylor, G., Porter, E., & Muñoz, R. V. (2022). Dataset of british english speech recordings for psychoacoustics and speech processing research : the clarity speech corpus. *Data in brief*, 41, 107951. <https://doi.org/10.1016/j.dib.2022.107951>
- Grange, J. A., & Culling, J. F. (2018). The factor analysis of speech : limitations and opportunities for cochlear implants. *Acta acustica united with acustica*, 104(5), 835-838. <https://doi.org/10.3813/aaa.919253>
- Smith, E. C., & Lewicki, M. S. (2006). Efficient auditory coding. *Nature*, 439(7079), 978-982. <https://doi.org/10.1038/nature04485>
- Ueda, K., & Nakajima, Y. (2017). An acoustic key to eight languages/dialects : factor analyses of critical-band-filtered speech. *Scientific reports*, 7(1). <https://doi.org/10.1038/srep42468>



# Troubles neurologiques de la parole et de la voix



# Contribution à l'étude psychométrique du Voice Handicap Index auprès de patients parkinsoniens

Virginie ROLAND<sup>1</sup>  
Véronique DELVAUX<sup>1,2</sup>  
Bernard HARMEGNIES<sup>1</sup>  
Myriam PICCALUGA<sup>1</sup>  
Kathy HUET<sup>1</sup>

<sup>1</sup> Service de Métrologie et Sciences du Langage, Institut de Recherche en Sciences et Technologies du Langage, UMONS, Belgique

<sup>2</sup> Fonds National de la Recherche Scientifique, Belgique

Les troubles de la voix dans la maladie de Parkinson (MP) constituent un des marqueurs clés de l'évolution de l'état de santé des patients, ces troubles ayant des répercussions importantes sur leurs capacités communicationnelles, et par-delà, sur leur qualité de vie (Miller et al., 2008). Afin de quantifier les conséquences psychosociales des personnes souffrant de pathologies vocales et de leur assurer un meilleur suivi, Jacobson et al. (1997) ont élaboré le Voice Handicap Index (VHI), questionnaire subjectif d'auto-évaluation des troubles vocaux selon 3 aspects : physique (AP), fonctionnel (AF) et émotionnel (AE). Plusieurs études ont étudié les propriétés psychométriques de cet outil auprès de patients MP à différents stades de la maladie dans différentes langues (Guimarães et al., 2017 ; Chen et al., 2022) mais aucune auprès d'une population MP francophone à un stade avancé de la maladie.

Notre étude a été menée auprès de 80 patients MP francophones de Belgique (44 hommes/36 femmes,  $\text{âge}_{\text{moyen}} = 71$  ans) à différents stades de la maladie (léger=19, modéré=49 et sévère=12). Ils ont été soumis à différents questionnaires/échelles pour documenter leur fonctionnement général (BDI-II (Beck, Steer & Brown, 1996), PDQ-39 (Auquier et al., 2002), MoCA (Nasreddine et al., 2005), degré de sévérité de la dysarthrie-DSD). Afin de mener l'analyse des caractéristiques psychométriques globales du VHI (fiabilité et validité), deux échelles d'auto-évaluation visuelles analogiques ont été utilisées ( $\text{VAS}_{\text{Voix}}$  et  $\text{VAS}_{\text{Communication}}$ ) ainsi que deux outils d'évaluation des interactions communicationnelles (SAID - Self Assessment of Interaction Disorders) (Pace et al., 2005) ; EVT - Evaluation Vocale par un Tiers (Dembinski & Yon, 2012)). Une analyse qualité des items (difficulté, discriminativité, fiabilité, validité) a également été menée.

La fiabilité test-retest et la fiabilité par consistance interne ont été évaluées grâce à la mesure de la corrélation ( $r_{\text{BP}}$ ) entre les scores (score total et aux trois dimensions AP, AF et AE) obtenus suite à deux passations du VHI à 7 jours d'intervalle, et à l'Alpha de Cronbach généralisé. La validité de construit a été évaluée sur base de corrélations avec d'autres outils (SAID, EVT,  $\text{VAS}_v$ ,  $\text{VAS}_c$ ) et par une Analyse en Composantes Principales (ACP).

Les résultats indiquent une fiabilité importante quelle que soit la méthode ( $r_{\text{BP}} \geq 0.71$ ,  $\text{alpha}_{\text{Cronbach}} \geq 0.83$ ), ainsi qu'une très bonne validité de construit ( $r_{\text{BP}} \geq 0.70$ ), à nuancer suivant les 3 dimensions. L'ACP met en évidence l'existence d'un facteur prédominant (36,62% de variance expliquée), les facteurs suivants portant moins d'informations (9 à 3% de variance expliquée), suggérant qu'une version abrégée du VHI pourrait être envisagée auprès de cette population. L'analyse des items suggèrent que des précisions pourraient être apportées dans leur formulation.

Par ailleurs, une analyse du lien entre le DSD et le VHI montre que ce dernier est un outil important d'évaluation et d'indication de prise en charge des problèmes vocaux rencontrés par les MP, même à des DSD légers.

### Références bibliographiques

- Auquier, P., Sapin, C., Ziegler, M. L., Tison, F., Destée, A., Dubois, B., Allicar, M. P., Thibault, J. L., Jenkinson, C., & Peto, V. (2002). Validation of the French language version of the Parkinson's Disease Questionnaire - PDQ-39. *PubMed*, 158(1), 41-50. <https://pubmed.ncbi.nlm.nih.gov/11938321>
- Beck, A. T., Steer, R. A., & Brown, G. (1996). Beck Depression Inventory–II (BDI-II), *APA PsycTests*. doi : 10.1037/t00742-000
- Chen, C., Lin, C., Su, C., Cheng, H., Chen, L., & Cherng, R. J. (2022). Evaluation of the Psychometric Properties of a Newly Developed Chinese Screening Tool for Speech Disorders in Patients With Parkinson's Disease. *Evaluation & the Health Professions*, 46(2), 127-134. <https://doi.org/10.1177/01632787221108458>
- Dembinski, C., & Yon, S. (2012). *La dysarthrie parkinsonienne : impact d'une rééducation vocale intensive sur le ressenti du handicap vocal par le couple lorsqu'un des conjoints présente des difficultés d'expression*. Université Claude Bernard.
- Guimarães, I., Cardoso, R. M., Pinto, S., & Ferreira, J. M. (2017). The Psychometric Properties of the Voice Handicap Index in People With Parkinson's Disease. *Journal of Voice*, 31(2), 258.e13-258.e18. <https://doi.org/10.1016/j.jvoice.2016.05.017>
- Jacobson, B. H., Johnson, A. F., Grywalski, C., Silbergleit, A. K., Jacobson, G. P., Benninger, M. S., & Newman, C. W. (1997). The Voice Handicap Index (VHI). *American Journal of Speech-language Pathology*, 6(3), 66-70. <https://doi.org/10.1044/1058-0360.0603.66>
- Miller, N., Noble, E., Jones, D., Allcock, L. M., & Burn, D. J. (2008). How do I sound to me ? Perceived changes in communication in Parkinson's disease. *Clinical Rehabilitation*, 22(1), 14-22. <https://doi.org/10.1177/0269215507079096>
- Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., Cummings, J. L., & Chertkow, H. (2005). The Montreal Cognitive Assessment, MoCA : A Brief Screening Tool For Mild Cognitive Impairment. *Journal of the American Geriatrics Society*, 53(4), 695-699. <https://doi.org/10.1111/j.1532-5415.2005.53221.x>
- Pace, C., Robert, D., Loundou, A., Azulay, J., Witjas, A., & Auquier, P. (2005). Évaluation du handicap communicationnel dans la maladie de Parkinson : développement et prévalidation d'une échelle d'auto-évaluation des troubles communicationnels. *Rééducation orthophonique*, 224, 81-100.

# Diagnostic différentiel des troubles moteurs de la parole grâce aux mouvements oro-moteurs et aux modulations de la parole

Marion BOURQUI<sup>1</sup>  
Monica LANCHEROS<sup>1</sup>  
Marina LAGANARO<sup>1</sup>

<sup>1</sup>Laboratoire de Neuropsycholinguistique,  
Faculté de Psychologie et des Sciences de l'Éducation,  
Université de Genève

Les troubles moteurs de la parole comprennent l'apraxie de la parole (AoS) et la dysarthrie. Ces deux troubles présentent des caractéristiques communes qui rendent le diagnostic différentiel parfois compliqué, tant sur le plan perceptif (Pernon et al., 2022) que dans les classifications basées sur les scores de déviance (Fougeron et al., 2022). Plusieurs tâches sont fréquemment utilisées pour évaluer ces deux types de troubles moteurs de la parole, mais aucune d'entre elles ne permet de distinguer ces deux pathologies de manière fiable. Par conséquent, dans cette étude, nous avons utilisé des tâches moins usuelles dans l'évaluation des troubles moteurs de la parole afin d'identifier si elles permettent de différencier l'AoS de la dysarthrie. Nous avons choisi des tâches qui ciblent spécifiquement les difficultés rencontrées dans les dysarthries : la production de mouvements oro-moteurs (Ziegler, 2002) et les modulations de la parole (Duffy, 2013).

Nous avons recruté 16 participants présentant des troubles moteurs de la parole : 8 participants atteints d'une AoS suite à une lésion acquise de l'hémisphère gauche et 8 participants atteints d'une dysarthrie hypokinétique consécutive à une maladie de Parkinson, les deux groupes étant appariés sur le score perceptif de la BECD (Auzou & Rolland-Monnoury, 2019). Ils ont tous effectué une tâche de production de mots, pseudo-mots, et mouvements oro-moteurs sonores appariés au niveau des mouvements articulatoires initiaux, ainsi qu'une tâche de production de pseudo-mots en parole normale versus chuchotée.

L'exactitude des réponses a été codée par deux juges indépendants, avec un haut niveau d'accord inter-juges. Concernant la tâche de production de séquences de parole et de mouvement oro-faciaux appariés, nous retrouvons un effet principal du type de stimuli (parole versus non-parole) et une interaction tendancielle entre le type de stimuli et la pathologie. Les contrastes indiquent une différence significative entre la production de parole et non-parole uniquement pour les participants atteints d'une AoS. Pour la tâche de modulations, nous retrouvons un effet principal du mode parole (normal versus chuchoté) avec des performances significativement meilleures en parole normale chez les deux groupes de participants.

Ces résultats préliminaires semblent indiquer qu'une tâche de parole versus non-parole permet de différencier les participants présentant une AoS versus une dysarthrie, ce qui semble rejoindre l'idée d'une apraxie orale souvent associée dans cette population (Whiteside et al., 2015). La tâche de modulations de la parole ne permet étonnement pas de cibler spécifiquement une des deux pathologies.

## Références bibliographiques

- Auzou, P., & Rolland-Monnoury, V. (2019). BECD: Batterie D'évaluation Clinique de la Dysarthrie [Clinical Dysarthria Assessment Battery].
- Duffy, J. R. (2013). *Motor speech disorders: Substrates, differential diagnosis, and management* (3rd ed.). Elsevier Mosby.
- Fougeron, C., Kodrasi, I., & Laganaro, M. (2022). Differentiation of Motor Speech Disorders through the Seven Deviance Scores from MonPaGe-2.0. s. *Brain Sciences*, 12(11), 1471.
- Pernon, M., Assal, F., Kodrasi, I., & Laganaro, M. (2022). Perceptual classification of motor speech disorders: the role of severity, speech task, and listener's expertise. *Journal of Speech, Language, and Hearing Research*, 65(8), 2727-2747.
- Whiteside, S. P., Dyson, L., Cowell, P. E., & Varley, R. A. (2015). The relationship between apraxia of speech and oral apraxia: association or dissociation?. *Archives of Clinical Neuropsychology*, 30(7), 670-682.
- Ziegler, W. (2002). Task-related factors in oral motor control: Speech and oral diadochokinesis in dysarthria and apraxia of speech. *Brain and language*, 80(3), 556-575.

# La désorganisation temporelle dans la dysarthrie ataxique : une altération au niveau global ou local

Angéline BOURBON<sup>1</sup>  
Cécile FOUGERON<sup>1</sup>  
Lise CREVIER-BUCHMAN<sup>1</sup>

<sup>1</sup> Laboratoire de Phonétique et Phonologie,  
UMR 7018,  
Université Sorbonne Nouvelle

La dysarthrie ataxique présente une atteinte multidimensionnelle de la parole. La caractéristique principale est un ralentissement du débit (Duffy, 2013), souvent accompagné d'une production de syllabes isochrones (e.g. Ackermann et Hertrich, 1994). Ces corrélats acoustiques principaux de la 'parole scandée', sont caractéristiques de ce trouble (Hartelius et al., 2000). Cependant, il a été montré que de tels déficits de l'organisation temporelle pourraient dépendre de la tâche de parole (Ziegler, 2002 ; Ziegler & Wessel, 1996) avec des résultats pouvant être contradictoires entre les études. En effet, une altération de l'organisation temporelle de la parole peut se manifester différemment selon le contenu linguistique (syllabes, phrases) (Staiger et al., 2017) et les instructions de performance (Maas, 2016).

Cette étude explore la capacité à contrôler l'organisation et la stabilité temporelle, en fonction de la naturalité et taille de l'item à répéter et l'instruction de débit. Nous avons enregistré 35 locuteurs francophones porteurs d'une dysarthrie ataxique, divisés en 4 groupes de sévérité, comparés à des locuteurs neurotypiques. Une tâche de répétition en continu pendant 15 secondes a été réalisée dans deux conditions d'instruction de débit (maximale vs. confortable). Les items à répéter étaient : des syllabes uniques (/ba/ et /go/), une séquence de syllabes (/badego/) et deux courtes phrases. Nous avons mesuré acoustiquement (i) la stabilité temporelle globale via un index temporel pris sur les 15 secondes, calculant les différences de durées normalisées entre deux items consécutifs (syllabes ou phrases, pauses exclues) tel que  $(| \text{Item } i - \text{Item } i-1 | / ((\text{Item } i + \text{Item } i-1) / 2))$  ; (ii) la stabilité temporelle locale, comparant la durée (en ms.) des syllabes dans des mots lexicaux ou dans la séquence trisyllabique.

La stabilité globale diffère en fonction de l'item répété : les locuteurs dysarthriques sont plus stables (index plus faible) en répétition de syllabes que de séquences ou de phrases, comme les locuteurs contrôles. Ce pattern est inattendu, la littérature indiquant davantage d'instabilité temporelle sur les tâches de production moins naturelles (e.g. Ziegler et Wessel, 1996). La stabilité locale montre que les locuteurs ataxiques parviennent à produire des contrastes de durées des syllabes dans les mots de la phrase et dans la séquence trisyllabique (i.e. parole moins naturelle) dans les deux conditions de débit. Ce résultat est et également inattendu (e.g. Hartelius et al. 2000). Ces résultats montrent que les locuteurs ataxiques maintiennent une organisation temporelle typique, même dans une parole moins naturelle, sollicitant davantage l'exécution motrice.

## Références bibliographiques

- Ackermann, H., & Hertrich, I. (1994). Speech rate and rhythm in cerebellar dysarthria: an acoustic analysis of syllabic timing. *Folia Phoniatrica et Logopaedica*, *46*, 70–78.
- Duffy, J. R. (2013). *Motor speech disorders: Substrates, differential diagnosis, and management*. St. Louis, MO: Elsevier Health Sciences.
- Hartelius, L., Runmaker, B., Andersen, O., & Nord, L. (2000). Temporal Speech Characteristics of Individuals with Multiple Sclerosis and Ataxic Dysarthria: ‘Scanning Speech’ Revisited. *Folia Phoniatrica et Logopaedica*, *52*, 228–238.
- Maas, E. (2016). Speech and nonspeech: What are we talking about? *International Journal of Speech-Language Pathology*, *19*, 345–359.
- Staiger, A., Schölderle, T., Brendel, B., Bötzel, K., & Ziegler, W. (2017). Oral Motor Abilities Are Task Dependent: A Factor Analytic Approach to Performance Rate. *Journal of Motor Behavior*, *49*(5), 482–493.
- Ziegler, W., & Weseel, K. (1996). Speech timing in ataxic disorders: Sentence production and rapid repetitive articulation. *Neurology*, *47*, 208–214.
- Ziegler, W. (2002). Task-related factors in oral motor control: Speech and oral diadochokinesis in dysarthria and apraxia of speech. *Brain and Language*, *80*, 556–575.

# Exploration des liens entre fluence, disfluences, organisation temporelle de la parole et performances cognitivo-langagières dans les aphasies primaires progressives

Fanny GUITARD-IVENT<sup>1</sup>  
Mélanie JUCLA<sup>1</sup>  
Corine ASTESANO<sup>1</sup>  
Jérémie PARIENTE<sup>2,3</sup>

<sup>1</sup>Laboratoire de NeuroPsychoLinguistique,  
Université de Toulouse, Toulouse, France

<sup>2</sup>ToNIC, Toulouse NeuroImaging Center,  
Université de Toulouse, Inserm, UPS, France

<sup>3</sup>Département de Neurologie, Pole Neurosciences, Hôpital Universitaire de Toulouse,  
Toulouse, France

L'étude du décours temporel de la parole et des phénomènes de disfluences font l'objet d'une attention grandissante dans l'appréhension des atteintes cognitivo-langagières résultant de maladies neurodégénératives comme la maladie d'Alzheimer ou les démences fronto-temporales (Pistono, Pariente & Jucla, 2022; Pistono et al., 2019). Dans les Aphasies Primaires Progressives, cela s'avère pertinent pour mieux caractériser le profil d'atteinte linguistique (Hoffman et al., 2017). Les APPnf (non fluentes) sont caractérisées par un débit ralenti du fait de nombreuses pauses vides. Les APPvs (sémantique) et APPlp (logopénique) sont dites fluentes. Néanmoins, un débit préservé peut cacher une organisation temporelle perturbée avec de nombreuses disfluences.

L'originalité de ce travail exploratoire est de 1) proposer une analyse fine de ces phénomènes à partir d'un plus grand spectre de variables dans un corpus de discours narratif sur support imagé obtenu auprès d'un échantillon de 13 locuteurs (3 APPnf, 4 APPlp, 3 APPvs, 3 contrôles, âge moyen 67 ans, min=56; max=75) ; 2) les relier aux performances des locuteurs à des épreuves standardisées d'évaluation des compétences lexicales.

Nous avons annoté manuellement les disfluences au sein des unités interpausales (IPU) – pauses silencieuses  $\geq 200$  ms (Bertrand et al., 2008) – sur PRAAT et extrait les variables présentées *Tableau 1*. Nous observons une grande variabilité entre locuteurs en termes de décours temporel du discours tant sur le temps de locution que sur la répartition en pauses silencieuses, disfluences sonores et parole fluente (voir *Figure 1a*). L'analyse détaillée de la durée relative de chaque type de disfluences sonores par rapport au temps parlé met en évidence l'importance temporelle des pauses pleines (*Figure 1b*). Les répétitions et les reformulations, moins fréquentes, occupent un temps de parole considérable chez certains locuteurs (ex. LOGO\_4, CTRL\_3). Si aucun groupe ne se distingue sur les aspects temporels et les disfluences, des relations intéressantes entre ces mesures et les performances cognitivo-langagières des locuteurs semblent émerger. En effet, les corrélations, tous groupes confondus, montrent un lien fort entre variables temporelles et processus d'accès lexical (durée d'IPU  $\sim$  Fluence,  $r=.71$ ), alors que les disfluences apparaissent plutôt liées à l'altération du système lexico-sémantique (Répétition/Reformulation  $\sim$  CompréhensionOrale,  $r=-.80$ ).

Cette étude nous a permis d'identifier des variables pertinentes jusque-là inexplorées (ex. durées d'IPU, durée relative des disfluences) qui permettront, sur un corpus plus étendu, combinées à d'autres paramètres, de mener des analyses en cluster afin de mieux comprendre le lien entre parole fluente, parole disfluente et atteinte cognitivo-langagière dans les APP.

Tableau 1. Variables analysées

Pauses silencieuses	Nombre d'occurrence, fréquence d'occurrence, durée moyenne
Autres disfluences	Pauses pleines, Allongements, Reformulations, Répétitions : nombre d'occurrences et durée relative sur temps articulé
Variables temporelles	Durée d'IPU, nombre de syllabe par IPU, ratio temps articulé/temps de locution, vitesse articulatoire, temps et taux de parole fluente (parole articulée hors disfluences)
Variables cognitivo-langagières	Performances aux épreuves d'appariement mot-image, dénomination d'images, fluences verbales (fruits, verbes et mots en V) (Bézy, Renard & Pariente 2016). Empan mnésique auditivo-verbal (chiffres).

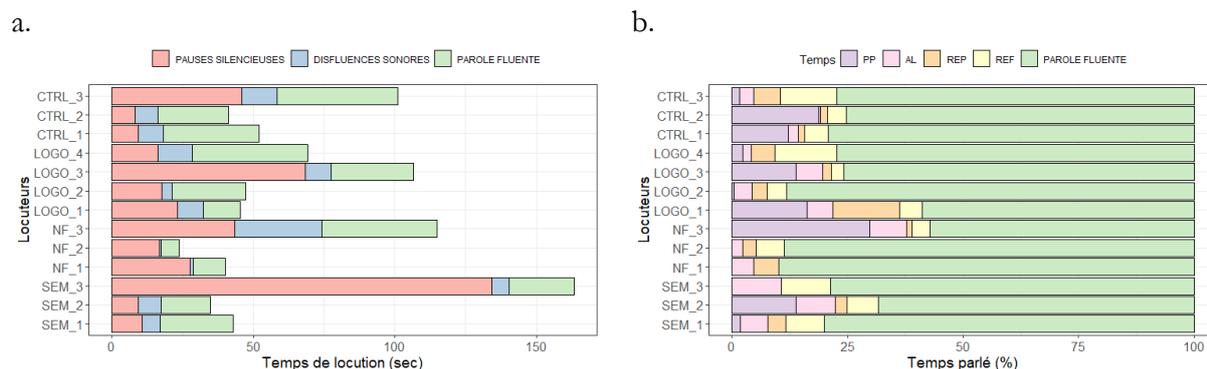


Figure 1. a) Répartition par locuteur (CTRL = contrôles, LOGO = APP logopénique, NF = APP non fluente, SEM = APP sémantique) du temps de locution (en secondes) en pause silencieuse, disfluences sonores et parole fluente (PF) ; b) Répartition détaillée par locuteur des disfluences sonores : pauses pleines (PP), allongements (AL), répétitions (REP), reformulation (REF) par rapport au % de temps parlé (hors pause silencieuse).

### Références bibliographiques

- Bertrand, R., Blache, P., Espesser, R., Ferré, G., Meunier, C., Priego-Valverde, B., & Rauzy, S. (2008). Le CID-Corpus of Interactional Data-Annotation et exploitation multimodale de parole conversationnelle. *Revue TAL*, 49(3), 105-134
- Bézy, C., Renard, A., & Pariente, J. (2016). *GRÉMOTS: évaluation du langage dans les pathologies neurodégénératives*.

- Hoffman, P., Sajjadi, S. A., Patterson, K., & Nestor, P. J. (2017). Data-driven classification of patients with primary progressive aphasia. *Brain and Language*, 174, 86-93. DOI 10.1016/j.bandl.2017.08.001.
- Pistono, A., Pariente, J., Bézy, C., Lemesle, B., Le Men, J. & Jucla, M. (2019). What happens when nothing happens? An investigation of pauses as a compensatory mechanism in early Alzheimer's disease. *Neuropsychologia*, 124, 133-143. DOI 10.1016/j.neuropsychologia.2018.12.018.
- Pistono, A., Pariente, J., & Jucla, M. (2022). Disfluency patterns in Alzheimer's disease and frontotemporal lobar degeneration. *Clinical Linguistics & Phonetics*, 1-14. DOI 10.1080/02699206.2022.2112085.



# Prédiction de la fluence à partir de mesures temporelles automatiques : vers une application clinique pour améliorer le diagnostic et le suivi des patients aphasiques

Lionel FONTAN<sup>1</sup>  
Typhanie PRINCE<sup>2</sup>  
Aleksandra NOWAKOWSKA<sup>2</sup>  
Halima SAHRAOUI<sup>3</sup>  
Silvia MARTÍNEZ-FERREIRO<sup>4</sup>

<sup>1</sup>Archean LABS

<sup>2</sup>Praxiling UMR 5267,  
Université Paul Valéry Montpellier 3

<sup>3</sup>Laboratoire de NeuroPsychoLinguistique E.A. 4156,  
Université de Toulouse

<sup>4</sup>Gerontology & Geriatrics research Group,  
University of A Coruña

La dichotomie clinique traditionnelle opposant aphasie *fluente* et aphasie *non fluente* fait depuis longtemps l'objet de critiques, notamment du fait du manque de fiabilité des jugements subjectifs effectués à cet égard (Gordon, 1998 ; Gordon & Clough, 2022). Par conséquent, le développement d'outils objectifs d'analyse de la fluence représente un enjeu majeur pour améliorer le diagnostic et le suivi clinique des patients aphasiques.

Dans cette optique, nous testons la pertinence d'un algorithme de mesure automatique de la fluence initialement développé dans le domaine de l'acquisition des langues (Fontan *et al.*, 2018 ; Fontan *et al.*, 2022 ; Detey *et al.*, 2020). Suivant le protocole élaboré dans le cadre du projet AADI (Sahraoui, Martínez-Ferreiro & Nowakowska, 2022), un recueil de données orales a été réalisé auprès de participants aphasiques et contrôles dans des tâches diversifiées, parmi lesquelles une épreuve de lecture de phrases issue du BDAE-HDAE (Goodglass *et al.*, 2001 ; Mazaux & Orgogozo, 1982). Les trois phrases les plus complexes de l'épreuve, produites par 34 participants (29 aphasiques, 5 contrôles), ont fait l'objet d'une évaluation par 3 orthophonistes sur le plan de la fluence au moyen d'une échelle qualitative (score de 1 à 5). Les accords inter-évaluateurs étant élevés (alpha de Cronbach = 0,95, rhos de Spearman  $\geq 0,86$ ), les scores de fluence ont été moyennés pour chacun des enregistrements.

Afin de prédire ces scores de référence, un algorithme de segmentation temporelle du signal (André-Obrecht, 1988) a été appliqué sur les enregistrements, et les segments obtenus ont été regroupés automatiquement en pseudo-syllabes (Farinas & Pellegrino, 2001) et en pauses silencieuses. Quatre prédicteurs ont été calculés : le ratio de parole (durée de parole/durée totale de l'enregistrement), le débit de pseudo-syllabes, l'écart-type de la durée des pseudo-syllabes, et le débit de pauses silencieuses. Les prédicteurs ont enfin été intégrés dans des modèles linéaires

(régression linéaire multiple) et non linéaires (forêt d'arbres décisionnels et machine à vecteurs de support), en appliquant une procédure de validation croisée.

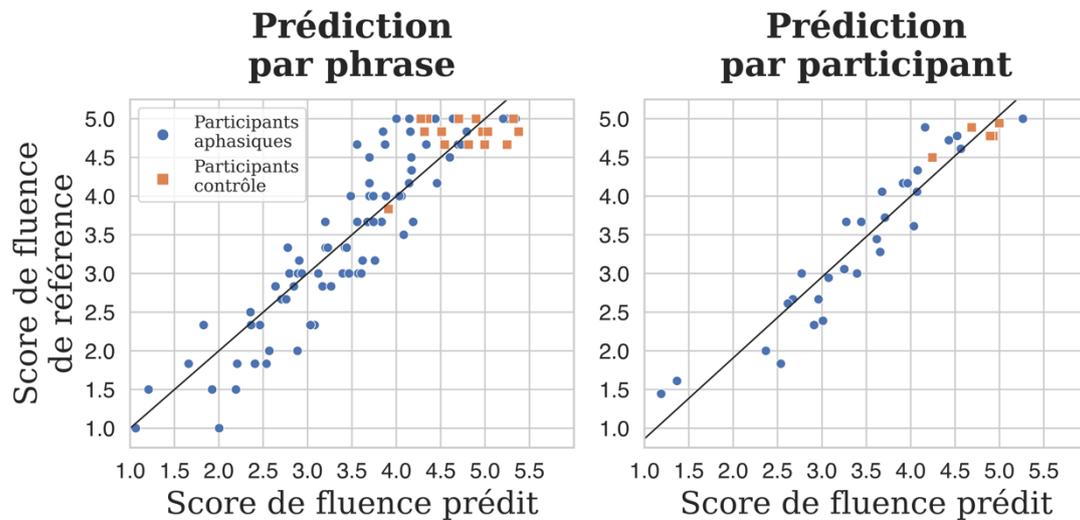


Figure 1. Prédiction des scores subjectifs de fluence en utilisant une régression linéaire multiple et l'ensemble des prédicteurs automatiques.

Les résultats montrent une précision élevée des prédictions, quel que soit le modèle. La régression linéaire, dont les paramètres sont les plus facilement interprétables, permet d'obtenir un coefficient de corrélation de 0,87 avec les scores de fluence attribués pour chaque phrase, et de 0,93 en agrégeant les données par participant. L'intégration d'un dernier paramètre automatique sensible aux répétitions et faux-départs permet d'améliorer significativement les prédictions, avec un coefficient de corrélation de 0,92 au niveau de la phrase, et de 0,95 au niveau du participant (Figure 1).

En conclusion, ces résultats prometteurs nous permettent d'envisager une application clinique de l'outil pour cette épreuve spécifique de lecture de phrases, et ouvrent des perspectives intéressantes et immédiates pour l'évaluation automatique de la fluence en discours spontané.

### Références bibliographiques

- André-Obrecht, R. (1988). A new statistical approach for the automatic segmentation of continuous speech signals. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 36(1), 29-40.
- Detey, S., Fontan, L., Le Coz, M., & Jmel, S. (2020). Computer-assisted assessment of phonetic fluency in a second language: a longitudinal study of Japanese learners of French. *Speech Communication*, 125, 69-79.
- Farinas, J., & Pellegrino, F. (2001). Automatic rhythm modeling for language identification. In *Proceedings of 7<sup>th</sup> European Conference on Speech Communication and Technology (EUROSPEECH)*, ISCA: 2539-2542.
- Fontan, L., Kim, J., De Fino, V., & Detey, S. (2022). Predicting speech fluency in children using automatic acoustic features. In *Proceedings of the 2022 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*, IEEE: 1085-1090.

- Fontan, L., Le Coz, M., & Detey, S. (2018). Automatically measuring L2 speech fluency without the need of ASR: A proof-of-concept study with Japanese learners of French. In *Proceedings of the Nineteenth Annual Conference of the International Speech Communication Association (INTERSPEECH)*, ISCA: 2544-2548).
- Goodglass, H., Kaplan, E., & Barresi, B. (2001). *Boston diagnostic aphasia examination*. Williams & Wilkins.
- Gordon, J. K. (1998). The fluency dimension in aphasia. *Aphasiology*, 12(7-8), 673-688.
- Gordon, J. K., & Clough, S. (2022). How do clinicians judge fluency in aphasia? *Journal of Speech, Language, and Hearing Research*, 65(4), 1521-1542.
- Mazaux, J.-M., & Orgogozo, J.-M. (1982). *Échelle d'évaluation de l'aphasie HDAE (adaptation française du Boston Diagnosis Aphasia Examination, H. Goodglass & A. Kaplan, 1972)*. Éditions Scientifiques et Psychologiques.
- Sahraoui, H., Martínez-Ferreiro, S., & Nowakowska, A. (2022). *Projet AADI : Aphasie et Analyse du Discours en Interactions : constitution de bases de données et nouvelles méthodes d'exploitation. Manuel d'instructions : méthodologie de recueil, transcription et codage des données*. Université de Toulouse Jean-Jaurès & Université Paul Valéry Montpellier 3. Disponible en ligne à l'adresse <https://hal-cnrs.archives-ouvertes.fr/hal-03913435>



# Prosodie



# Impact de la maladie de Parkinson sur les paramètres acoustiques de la prosodie : une revue systématique et une méta-analyse Bayésienne

Jules FUMEL<sup>1</sup>  
Delphine BAHUAUD<sup>2</sup>  
Ethan WEED<sup>3,4</sup>  
Riccardo FUSAROLI<sup>3,4</sup>  
Anahita BASIRAT<sup>1</sup>

<sup>1</sup>Univ. Lille, CNRS, UMR 9193 SCALab, F-59000 Lille, France

<sup>2</sup>Département d'orthophonie, Faculté de Médecine, UFR3S, Université de Lille, France

<sup>3</sup>Linguistics, Cognitive Science and Semiotics, School of Communication and Culture, Aarhus University, Aarhus, Denmark

<sup>4</sup>Interacting Minds Center, School of Culture and Society, Aarhus University, Aarhus, Denmark

La dysarthrie hypokinétique de la maladie de Parkinson (MP) se caractérise par une insuffisance prosodique s'exprimant par une diminution de l'accentuation, une monotonie d'intensité et de hauteur, une variabilité du débit et une possible imprécision des phonèmes (Pinto et al., 2010). Si les caractéristiques générales de la dysarthrie hypokinétique sont à l'heure actuelle connues, l'ampleur des modifications de la parole ainsi que la quantification des effets induits par l'atteinte des ganglions de la base restent à déterminer (Van Lancker-Sidti et al., 2006). Ainsi, cette contribution a pour objectif de pouvoir examiner de manière fiable, en se basant sur des paramètres acoustiques et sous forme d'une méta-analyse bayésienne, l'effet de la MP sur l'expression de la prosodie.

Dans cette étude, nous avons comparé les scores d'individus atteints de la MP avec ceux de sujets contrôles sur différentes variables. Nous nous sommes concentrés sur la fréquence fondamentale et sa variabilité, l'intensité et sa variabilité, le débit et les pauses. Une recherche systématique a été effectuée dans *PubMed*, *PsycINFO* et *Web of science*. Après l'analyse des résultats de cette recherche, nous avons exclu les articles qui ne correspondaient pas aux critères d'inclusion. Les articles conservés présentaient des résultats pour une ou plusieurs tâches d'intérêt. Pour chaque combinaison de mesure acoustique et de tâche, les tailles d'effet (g de Hedges) ont été extraites en comparant les groupes d'individus atteints de la MP aux groupes contrôles. Nous avons estimé les tailles d'effet en utilisant des modèles de régression hiérarchique bayésienne (ex. Weed & Fusaroli, 2020).

La taille d'effet globale estimée concernant la variabilité de F0, calculée à partir de 47 tailles d'effet issues de 20 articles, était de -0,75 (IC 95% : [-1,01 ; -0,48] ; ratio de preuve : infini ; crédibilité : 100%). Celle concernant la durée des pauses, calculée à partir de 20 tailles d'effets issues de 9 articles, était de 0,49 (IC 95% : [0,01 ; 0,94] ; ratio de preuve : 41,11 ; crédibilité : 98%). Les tailles d'effet globales relevées pour les autres variables ne montraient pas de différence notable entre les deux groupes. Ces premiers résultats apportent des précisions sur l'implication des

ganglions de la base dans la production de la parole sur le plan prosodique ainsi que sur les caractéristiques de la parole hypokinétique. La comparaison avec certaines pathologies, telle l'ataxie (Ackermann & Hertrich, 2000), pourrait permettre de préciser les caractéristiques d'autres types de dysarthrie et contribuer à mieux comprendre les mécanismes neurocognitifs sous-jacents à la production de la parole, notamment le rôle des ganglions de la base et du cervelet.

### Références bibliographiques

- Ackermann, H., & Hertrich, I. (2000). The contribution of the cerebellum to speech processing. *Journal of Neurolinguistics*, 13(2-3), 95-116. [https://doi.org/10.1016/S0911-6044\(00\)00006-3](https://doi.org/10.1016/S0911-6044(00)00006-3)
- Pinto, S., Ghio, A., Teston, B., & Viallet, F. (2010). La dysarthrie au cours de la maladie de Parkinson. Histoire naturelle de ses composantes : Dysphonie, dysprosodie et dysarthrie. *Revue Neurologique*, 166 (10), 800-810. <https://doi.org/10.1016/j.neurol.2010.07.005>
- Van Lancker Sidtis, D., Pachana, N., Cummings, J. L., & Sidtis, J. J. (2006). Dysprosodic speech following basal ganglia insult : Toward a conceptual framework for the study of the cerebral representation of prosody. *Brain and Language*, 97(2), 135-153. <https://doi.org/10.1016/j.bandl.2005.09.001>
- Weed, E., & Fusaroli, R. (2020). Acoustic measures of prosody in right-hemisphere damage: A systematic review and meta-analysis. *Journal of Speech, Language, and Hearing Research*, 63(6), 1762-1775. [https://doi.org/10.1044/2020\\_JSLHR-19-00241](https://doi.org/10.1044/2020_JSLHR-19-00241)

# Caractérisation de la prosodie du mot chez des enfants avec autisme avec et sans déficit phonologique

Sandrine FERRÉ<sup>1</sup>  
Marlène GASNIER<sup>2</sup>  
Bénédicte GRANDON<sup>3</sup>

<sup>1</sup> UMR 1253, iBrain,  
Université de Tours, Inserm, Tours

<sup>2</sup> Université de Tours

<sup>3</sup> Universität Oldenburg & Hearing4All Cluster of Excellence

La prosodie des personnes avec autisme est souvent qualifiée d'atypique (e.g., Peppé et al., 2007). La variabilité des mesures utilisées (notamment de F0, mais aussi de l'intensité, ou de la durée, de la voyelle à l'énoncé), des outils et des sujets testés n'aboutit à aucun consensus solide quant à la caractérisation de cette déviation. Or les liens étroits qu'entretiennent les structures prosodiques et phonologiques suggèrent qu'une altération dans un domaine pourrait déséquilibrer l'ensemble. Notre étude vise à déterminer en quoi les aspects prosodiques sont altérés dans l'autisme en particulier en contexte de déficit phonologique.

Nous avons analysé les propriétés acoustiques des voyelles /a, i, u/ produites dans le cadre d'une tâche de répétition de non-mots (LITMUS-QU-NWR-FR, Ferré & dos Santos, 2015), contenant des items de 1 à 3 syllabes, par deux groupes de quinze enfants âgés de 6 à 12 ans, présentant, l'un, un trouble du spectre de l'autisme (TSA) avec (TSA<sub>pi</sub>) ou sans (TSA<sub>pn</sub>) déficit phonologique, et, l'autre, un développement typique (DT). Nous avons mesuré (avec Praat) puis comparé la F0 moyenne, l'étendue de F0 et la durée des voyelles de chaque item et chaque groupe.

La F0 moyenne est significativement plus élevée chez les enfants TSA<sub>pi</sub> que chez les enfants sans déficit phonologique (TSA<sub>pn</sub> et DT). Il en est de même pour l'étendue de F0 sur la 1<sup>ère</sup> voyelle. Concernant la durée, les enfants avec TSA présentent des voyelles plus longues que les enfants DT pour tous les non-mots. Pour les non-mots trisyllabiques, nous observons un allongement significatif de la 2<sup>e</sup> voyelle pour les enfants TSA<sub>pi</sub>, mais de la 3<sup>e</sup> voyelle pour les TSA<sub>pn</sub>.

Ces résultats attestent de l'importance de ne pas considérer les enfants avec TSA comme un groupe langagier unique (Silleresi et al., 2020 ; Eigsti & Shuh, 2017). Nos résultats sur la durée, retrouvés chez Diehl & Paul (2013) et Olivati et al. (2017), le confirment : en contexte de déficit phonologique, l'allongement n'est pas marqué sur la voyelle finale, laissant supposer une interaction forte entre les deux niveaux langagiers.

L'analyse des voyelles à travers le prisme du déficit phonologique dans l'autisme a permis d'éviter certains écueils des études sur la prosodie dans cette population. Pour autant, cette étude n'observe qu'une infime partie de ce qu'est la prosodie, il est nécessaire de continuer d'observer chacun des aspects de la structure afin de comprendre la complexité des mécanismes à l'œuvre.

## Références bibliographiques

- Diehl, J. J., & Paul, R. (2013). Acoustic and perceptual measurements of prosody production on the profiling elements of prosodic systems in children with autism spectrum disorders. *Applied Psycholinguistics*, *34*(1), 135-161.
- Eigsti, I.M., & Schuh, J.M. (2017). Language acquisition in ASD: Beyond standardized language measures. In L.R. NAIGLES. (ed.), *Innovative investigations of language in autism spectrum disorder*, 183–200. De Gruyter.
- Ferré, S., & Santos, C. D. (2015). Comment évaluer la phonologie des enfants bilingues ? *Lidil*, *51*, 11-34. <https://doi.org/10.4000/lidil.3678>
- Olivati, A. G., Assumpção, F. B., & Misquiatti, A. R. N. (2017). Acoustic analysis of speech intonation pattern of individuals with Autism Spectrum Disorders. In *CoDAS* (Vol. 29). Sociedade Brasileira de Fonoaudiologia.
- Peppé, S., Mccann, J., Gibbon, F., O'hare, A., & Rutherford, M. (2007). Receptive and Expressive Prosodic Ability in Children With High-Functioning Autism. *Journal of Speech, Language, and Hearing Research*, *50*(4), 1015–1028
- Silleresi, S., Prevost, P., Zebib, R., Bonnet-Brilhault, F., Conte, D., & Tuller, L. (2020). Identifying language and cognitive profiles in children with ASD via a cluster analysis exploration: Implications for the new ICD-11. *Autism Research*, *13*(7), 1155–1167.

# Variations prosodiques lors de la synchronisation de parole avec un rythme externe : comparaison de locuteurs qui bégaièrent et normofluents

Maëva GARNIER<sup>1</sup>  
Anneke SLIS<sup>1</sup>  
Pascal PERRIER<sup>1</sup>  
Christophe SAVARIAUX<sup>1</sup>

<sup>1</sup>Univ. Grenoble Alpes, CNRS, Grenoble INP, GIPSA-lab, 38000 Grenoble, France

Diverses conditions améliorent la fluence de personnes qui bégaièrent, comme le fait de suivre un autre patron rythmique : par exemple en parlant de façon isochrone (Brady 1969; Hanna & Morris 1977; Hutchinson & Navarre, 1977; Kalinowski et al., 2000), en chantant (Perkins et al., 1976; Healey et al. 1976; Glover et al. 1996; Falk et al. 2016) ou en suivant la voix d'autres personnes (Adams & Ramig, 1980; Kalinowski et al., 2000; Ingham et al. 2006).

L'objectif de cette étude est de mieux comprendre l'effet positif de ces stimulations rythmiques externes, et d'examiner l'hypothèse qu'elles viennent pallier chez les personnes qui bégaièrent des difficultés à séquencer leur parole en suivant un patron prosodique complexe.

Pour cela, nous avons comparé la production de mêmes énoncés de sept syllabes (e.g. "Pattie passa la pagaie") par des personnes qui bégaièrent (PQB, N=16) et des personnes normofluentes (PNF, N=16) dans six conditions expérimentales : naturellement, sans contrainte rythmique (REPEAT) ; en synchronisant chaque syllabe avec un métronome audio de 120 BPM (ISO\_SYNC\_120), de 240 BPM (ISO\_SYNC\_240), et avec deux séquences non isochrones (NONISO\_SYNC\_1 et NONISO\_SYNC\_2) ; en suivant une séquence rythmique aperiodique imprédictible (REACT). Nous avons mesuré l'intensité acoustique, la fréquence fondamentale ( $f_0$ , en demi-tons) et la durée de chaque syllabe.

Comme dans plusieurs études antérieures (Arbisi-Kelm, 2005; Bergmann, 1986; Franke et al., 2020), les PQB ont montré des différences significatives par rapport aux PNF dans la réalisation prosodique de leurs énoncés "naturels" (variations réduites de  $f_0$  et de durée syllabique ; variations d'intensité plus marquées), tout en respectant le patron prosodique du Français – caractérisé par des excursions de  $f_0$  et des allongements syllabiques en fin de groupes accentuels et intonatifs (Jun & Fougeron, 2000).

Chez les PQB comme les PNF, la parole produite en suivant un rythme externe montre des variations de la  $f_0$  et de l'intensité de leurs syllabes par rapport à la parole "naturelle", mais qui continuent néanmoins de suivre le même patron prosodique. Dans les deux groupes, la durée syllabique est davantage influencée par la stimulation externe : elle tend à devenir quasi isochrone pour un métronome rapide, et à suivre la durée des notes d'une séquence non-isochrone.

On note que certaines conditions (REACT, ISO\_SYNC\_240) contribuent à réduire les différences observées entre PQB et PNF, ce qui pourrait expliquer leur effet bénéfique sur la fluence des PQB. Cependant, une telle convergence entre PQB et PNF n'est pas observée dans la condition ISO\_SYNC\_120 où les disfluences sont pourtant les plus réduites.

Dans leur ensemble, ces différents résultats infirment l'hypothèse que les personnes qui bégaient présentent un trouble majeur de planification de leur prosodie, qu'une stimulation rythmique externe pallie en remplaçant le patron prosodique complexe par un patron rythmique plus simple.

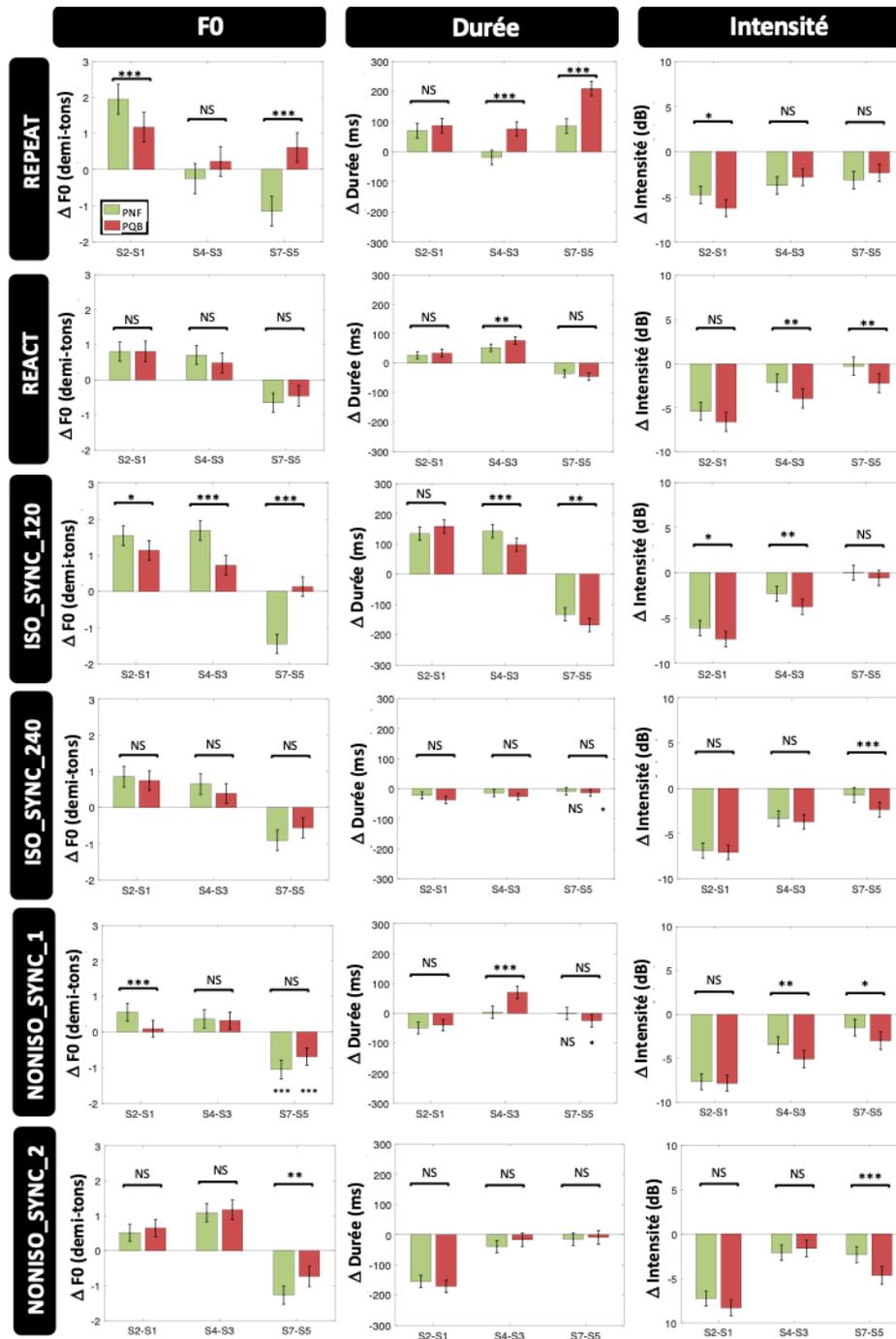


Figure 1. Variation d'intensité, de f0 et de durée en frontière de groupes accentuels (Syllabes 2, 4 et 7) pour des énoncés de 7 syllabes (e.g. « Pattie | passa | la pagaie ») produits par des personnes normofluentes (PNF) ou qui bégaient (PQB), sans contrainte rythmique (REPEAT), en suivant une stimulation rythmique a périodique non prédictible (REACT), en synchronisation avec un métronome de 120 BPM (ISO\_SYNC\_120) ou de 240 BPM (ISO\_SYNC\_240), ou avec deux rythmes non-isochrones (NONISO\_SYNC\_1, NONISO\_SYNC\_2).

## Références bibliographiques

- Arbisi-Kelm T. (2005). An Analysis of the Intonational Structure of Stuttered Speech. In *proc. Interspeech 2005, Lisbon*, 2405-2408.
- Bergmann, G. (1986). Studies in stuttering as a prosodic disturbance. *Journal of Speech and Hearing Research*, 29, 290-300.
- Brady, J. (1969). Studies on the metronome effect on stuttering. *Behaviour Research and Therapy*, 7(2), 197-204.
- Falk, S., Maslow, E., Thum, G., & Hoole, P. (2016). Temporal variability in sung productions of adolescents who stutter. *Journal of Communication Disorders*, 62, 101-114. <https://doi.org/10.1016/j.jcomdis.2016.05.012>
- Franke, M., Kello, C., Hoole, P., & Falk, S. (2020). Temporal organization during oral reading in children and adolescents who stutter. in *proc. International Seminar on Speech Production, Yale*.
- Glover, H., Kalinowski, J., Rastatter, M. P., & Stuart, A. M. (1996). Effect of Instruction to Sing on Stuttering Frequency at Normal and Fast Rates. *Perceptual and Motor Skills*, 83(2), 511-522. <https://doi.org/10.2466/pms.1996.83.2.511>
- Hanna, R., & Morris, S. (1977). Stuttering, speech rate, and the metronome effect. *Perceptual and Motor Skills*, 44(2) 452-454.
- Healey, E., Mallard A., & Adams, M. (1976). Factors contributing to the reduction of stuttering during singing. *Journal of Speech and Hearing Research*, 19(3), 475-480.
- Hutchinson, J., & Navarre, B. (1977). The effect of metronome pacing on selected aerodynamic patterns of stuttered speech: Some preliminary observations and interpretations. *Journal of Fluency Disorders*, 2(3), 189-204.
- Jun, S., & Fourgeron, C. (2000). A phonological model of French intonation. In : *Intonation*. Springer, Dordrecht, 209-242.
- Kalinowski, J., Stuart, A. M., Rastatter, M. P., Snyder, G. F., & Dayalu, V. N. (2000). Inducement of fluent speech in persons who stutter via visual choral speech. *Neuroscience Letters*, 281(2-3), 198-200. [https://doi.org/10.1016/s0304-3940\(00\)00850-8](https://doi.org/10.1016/s0304-3940(00)00850-8)
- Perkins, W. A., Rudas, J., Johnson, L. Y., & Bell, J. M. (1976). Stuttering : Discoordination of Phonation with Articulation and Respiration. *Journal of speech and hearing research*, 19(3), 509-522. <https://doi.org/10.1044/jshr.1903.509>



# Une nouvelle procédure psychoacoustique pour caractériser la perception de la prosodie après AVC de l'hémisphère droit

Aynaz ADL ZARRABI<sup>1</sup>  
Jean-Julien AUCOUTURIER<sup>1\*</sup>  
Lionel NACCACHE<sup>2\*</sup>  
Marie VILLAIN<sup>2</sup>

<sup>1</sup>FEMTO-ST Institute UMR 6174,  
CNRS/UBFC, Besançon

<sup>2</sup> Institut du Cerveau et de la Moëlle Epinière,  
APHP/Hôpital Pitié-Salpêtrière, Paris

\* *Les auteurs remercient le financement  
Appel à Projet Laboratoire 2021 de la Fondation Pour l'Audition*

86% des patients ayant eu un accident vasculaire cérébral (AVC) sont concernés par des déficits auditifs, et parmi ceux-ci, les patients d'AVC de l'hémisphère droit sont particulièrement sujets aux déficits de perception de la prosodie (Côté et al., 2007 ; Ponsot et al., 2018 ; Burred et al., 2019). Pourtant, les déficits de perception de la parole après AVC ne sont pas aussi bien étudiés que les autres symptômes « plus évidents » comme les symptômes verbaux ou moteurs, et les cliniciens ont besoin d'outils diagnostiques plus affinés et adaptés.

Dans cette étude, nous avons testé le potentiel d'une nouvelle méthode de psychophysique, la corrélation inverse (*reverse correlation* ; Ponsot et al., 2018), pour caractériser les déficits de perception de la prosodie interrogative chez des patients AVC de l'hémisphère droit. Le principe de la *reverse correlation* est de présenter aux patients un grand nombre de paires d'enregistrements d'un même mot (ex. le mot « vraiment ») dont les contours de hauteur (*pitch*) ont été manipulés aléatoirement par des méthodes informatiques (Burred et al., 2019), et de leur demander à chaque paire quelle est la prononciation la plus interrogative (« vraiment ! » vs « vraiment ? » - la deuxième). À partir des réponses à un grand nombre de paires aléatoires, la méthode produit plusieurs mesures: d'une part, un contour de *pitch* correspondant à la «représentation mentale» qu'a un patient d'une prosodie interrogative, et qui peut être comparé à la représentation typique d'un groupe contrôle; d'autre part, une mesure de la consistance, dite de «bruit interne» avec laquelle le patient applique cette représentation dans ces décisions.

Vingt-deux patients (âge moyen 57 ans, AVC droit sans symptômes aphasiques), ainsi que douze contrôles appariés, ont réalisé la tâche de jugement, et répondu à des questionnaires cliniques évaluant leur perception de la prosodie (MEC ; Joanette et al., 2004), l'attention auditive (AIRTAC ; Tessier & Weill-Chounlamounry, 2014) et les symptômes dépressifs et anxieux (HADS ; Zigmond et al., 1983).

Nos résultats montrent que la méthode a une bonne validité concourante: la représentation mentale du groupe patient diffère significativement de celle des contrôles et cette distance, ainsi que le bruit interne de patients, corrèle négativement avec les capacités prosodiques mesurées par la MEC et les capacités auditives mesurées par AIRTAC ; qu'elle présente plus de sensibilité que les méthodes existantes, en révélant des représentations mentales anormales pour des patients pourtant négatifs à la MEC ; et qu'elle permet d'établir de nouveaux profils cliniques de patients,

notamment d'identifier ceux qui ont des persévérations dans leurs réponses (bon bruit interne, mauvaises représentations).

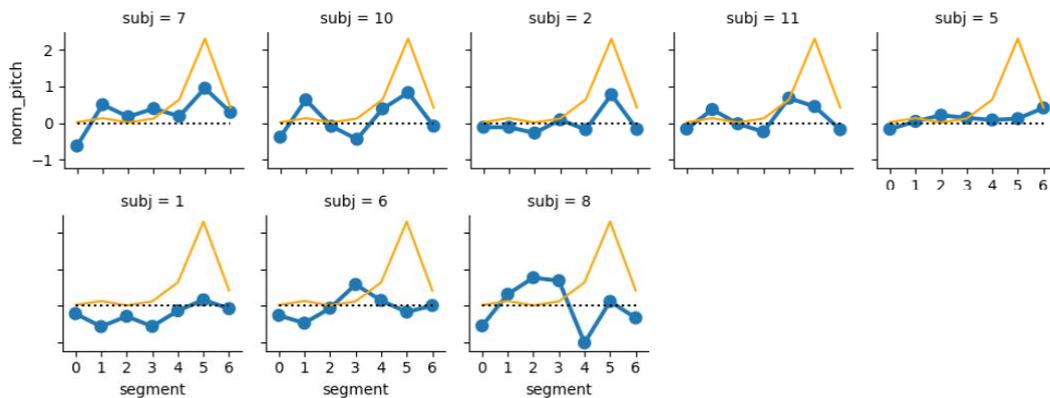


Figure 1. Exemples de représentations mentales de patients AVC (bleu) et d'un groupe contrôle apparié (orange) pour une tâche de détection de prosodie interrogative. La représentation contrôle (abscisses : temps, ordonnées : hauteur/*pitch*) fait apparaître un *pitch* croissant en fin de mot (*final pitch rise*) caractéristique de la prosodie interrogative en français. Les patients sont triés par distance croissante à cette représentation contrôle et, pour les plus graves, ne basent pas leurs décisions sur ce profil prosodique typique.

### Références bibliographiques

- Burred, J.J., Ponsot, E., Goupil, L., Liuni, M., & Aucouturier, J.J. (2019). Cleese: An open-source audio-transformation toolbox for data-driven experiments in speech and music cognition. *PLoS One*, 14(4), e0205943.
- Côté, H., Payer, M., Giroux, F., & Joanette, Y. (2007). Towards a description of clinical communication impairment profiles following right hemisphere damage. *Aphasiology*, 21(6-8), 739-749.
- Joanette, Y., Ska, B., & Côté, H. (2004). *Protocole Montréal d'Évaluation de la Communication (MEC)*. Isbergues: Ortho Ed. Published online.
- Ponsot, E., Burred, J.J., Belin, P., & Aucouturier, J.J. (2018). Cracking the social code of speech prosody using reverse correlation. *Proceedings of the National Academy of Sciences*, 115(15), 3972-3977.
- Tessier, C., & Weill-Chounlamounry, A. (2014). *Aide informatisée pour la rééducation des troubles auditifs centraux 2 (Airtac2)*. Paris Fr: Ortho Ed. Published online.
- Zigmond, A.S., & Snaith, R.P. (1983). The Hospital Anxiety and Depression Scale. *Acta Psychiatr Scand*, 67(6), 361-370.

# Représentation automatique du rythme de la parole pathologique via le spectre de modulations d'amplitude

Robin VAYSSE<sup>1,2</sup>  
Corine ASTESANO<sup>2</sup>  
Jérôme FARINAS<sup>1</sup>

<sup>1</sup>IRIT, Université de Toulouse, CNRS, Toulouse INP, UT3, Toulouse, France

<sup>2</sup>Laboratoire de Neuro-Psycho-Linguistique LNPL, Université Toulouse II, France

Une bonne structuration prosodique en groupements cohésifs hiérarchisés optimise la segmentation de la parole et l'accès au sens pour l'auditeur. Parmi les trois principes organisateurs de la parole (intonation, accentuation et rythme ; Di Cristo, 2011), le rythme constitue le socle de la prosodie puisqu'il permet l'intégration temporelle des proéminences accentuelles et des unités intonatives en lien avec les règles métriques sous-jacentes d'une langue donnée. Les processus de programmation de la parole (pathologique ou non) induisent des disfluences caractérisées par des ruptures du rythme de la parole, pouvant nuire à la compréhension et à l'intelligibilité de la parole. Caractériser le rythme linguistique est donc un enjeu fondamental en linguistique et en traitement automatique de la parole. Au-delà des méthodes basées sur le rythme linéaire (notamment pour la classification rythmique automatique des langues), il est nécessaire de développer des méthodes permettant de rendre compte de l'intégration temporelle de la hiérarchie prosodique (Arvaniti, 2012).

Notre recherche propose une représentation automatique du rythme à partir du signal de parole. Cette méthode permet de visualiser les répartitions d'énergie des différents niveaux prosodiques : syllabes, mot prosodique (pw ; Astésano, 2019), syntagme accentuel (AP), syntagme intonatif (IP ; Di Cristo, 2011). Le spectre de modulation d'amplitude (Envelope Modulation Spectrum – EMS) est obtenu en calculant l'enveloppe du signal auquel est appliqué un filtre 300-1000 Hz afin de capturer l'énergie des voyelles (Tilsen & Johnson, 2008 ; Vaysse et al., 2020 ; voir courbe noire, *Figure 1*). Un spectre de puissance est ensuite appliqué pour représenter les fréquences de 0 à 10 Hz (courbe orange). Un lissage (courbe rouge pointillée) est également appliqué pour englober les constituants prosodiques de niveau similaire. La détection automatique des constituants prosodiques a été validée par une annotation prosodique manuelle sur un échantillon du corpus (10 patients et 10 contrôles).

Le corpus de texte lu par 137 témoins, 86 patients atteints de cancers VADS et 205 patients atteints de la maladie de Parkinson est issu du projet RUGBI (Woisard et al., 2021). La *Figure 2* illustre la répartition spectrale rythmique sur les deux populations de patients.

La répartition énergétique sur l'EMS permet de visualiser les stratégies rythmiques des patients (préservation des cibles articulatoires au détriment de la fluence rythmique par une structuration essentiellement aux niveaux prosodiques inférieurs vs préservation de la fluence et des groupes prosodiques hiérarchisés). Nous présenterons les différents profils rythmiques caractérisant les pathologies et nous ouvrirons sur les applications potentielles pour l'évaluation automatique longitudinale des patients.

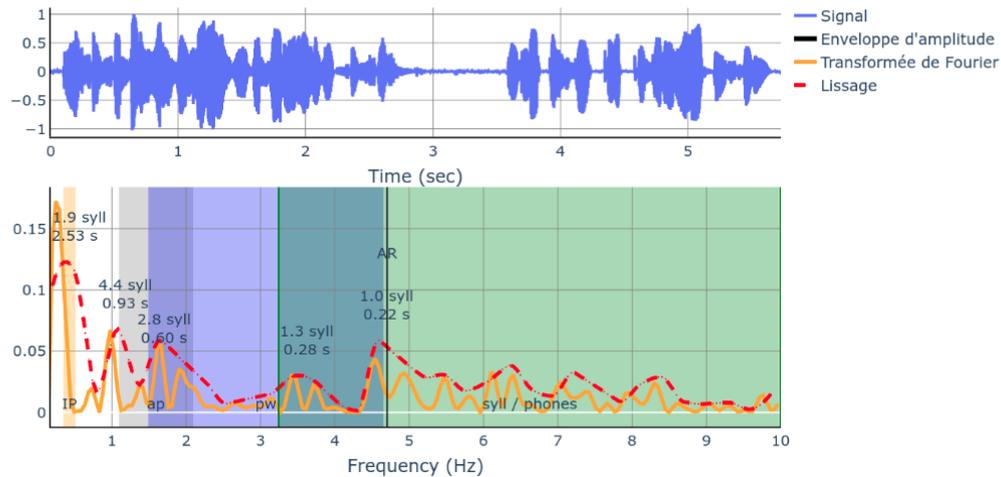


Figure 1 : EMS d'un locuteur contrôle sur l'extrait "Monsieur Seguin n'avait jamais eu de bonheur avec ses chèvres. Il les perdait toutes de la même façon". Les intervalles correspondant aux niveaux prosodiques annotés manuellement sont indiqués en couleur : orange pour l'IP, gris pour l'ap, bleu pour le pw et vert pour la syllabe. AR (*articulation rate*) désigne le débit articulatoire.

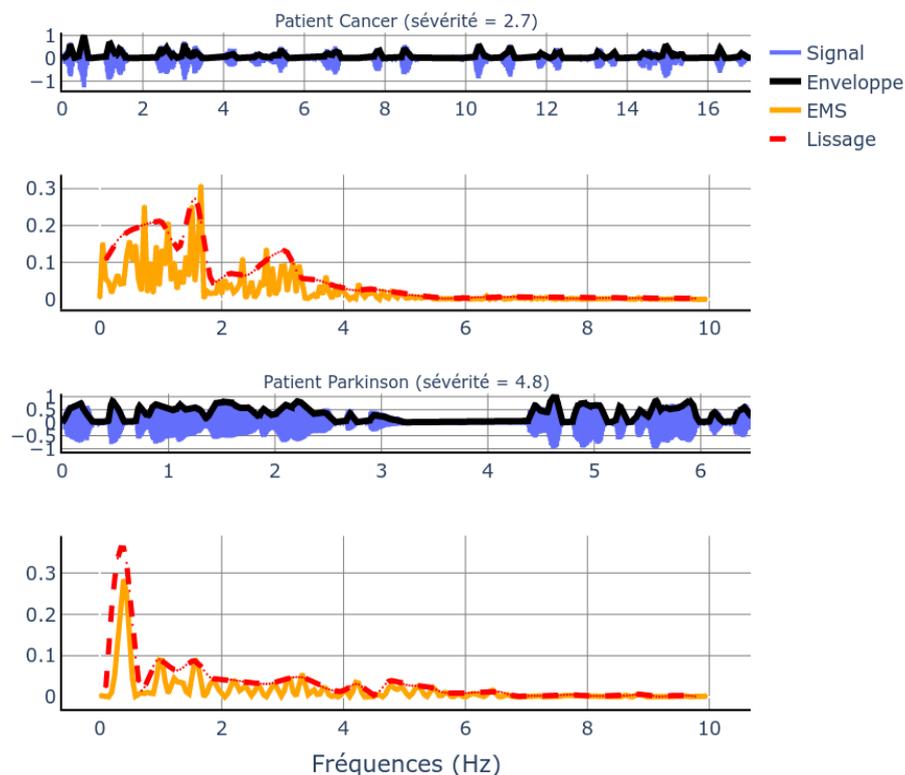


Figure 2 : Comparaison des EMS de deux patients sur la lecture des phrases "Monsieur Seguin n'avait jamais eu de bonheur avec ses chèvres. Il les perdait toutes de la même façon". L'EMS du sujet cancer VADS est en haut, celui du patient Parkinson est en bas. Les patients sont des personnes avec une sévérité de maladie élevée. L'énergie des spectres de modulations pour ces patients est plus concentrée vers les basses fréquences que pour le locuteur contrôle sur la

Figure 1.

## Références bibliographiques

- Arvaniti, A. (2012). The Usefulness of Metrics in the Quantification of Speech Rhythm. *Journal of Phonetics*, 3, 351–373.
- Astésano, C. (2019). The prosodic word as the domain of French accentuation - Empirical evidence. *Phonetics and Phonology in Europe, PaPE 2019*, Lecce : 170-171.
- Di Cristo, A. (2011). Une approche intégrative des relations de l'accentuation au phrasé prosodique du français. *Journal of French Language Studies*, 21(1), 73-95.
- Tilsen, S., & Johnson, K. (2008). Low-frequency fourier analysis of speech rhythm. *The Journal of the Acoustical Society of America*, 124(2), 34–39. <https://doi.org/10.1121/1.2947626>
- Vaysse, R., Farinas, J., Astésano, C., & André-Obrecht, R. (2021). Automatic Extraction of Speech Rhythm Descriptors for Speech Intelligibility Assessment in the Context of Head and Neck Cancers. *Interspeech 2021*, 1912-1916, <https://doi.org/10.21437/Interspeech.2021-1736>
- Woisard, V., Astésano, C., Balaguer, M., Farinas, J., Fredouille, C., Gaillard, P., Ghio, A., Giusti, L., Laaridh, I., Lalain, M., Lepage, B., Mauclair, J., Nocaudie, O., Piquier, J., Pouchoulin, G., Puech, M., Robert, D., & Roger, V. (2021b). C2SI corpus : a database of speech disorder productions to assess intelligibility and quality of life in head and neck cancers. *Language Resources and Evaluation*, 55(1), 173-190. <https://doi.org/10.1007/s10579-020-09496-3>



# Evaluation de la parole : vers la clinique



# Effet de la tâche sur la voix et la parole de patientes dysphoniques avec suspicion d'atteinte du nerf laryngé supérieur

Claire PILLOT-LOISEAU<sup>1</sup>  
Fabien BOURDIN<sup>1</sup>  
Jeevya AROUN<sup>1</sup>  
Manon HENNECHART<sup>1</sup>  
Diane LAZARD<sup>2</sup>  
Christophe TRESALLET<sup>3</sup>

<sup>1</sup>Laboratoire de Phonétique et Phonologie, UMR 7018 CNRS, Université Sorbonne Nouvelle, Paris, France

<sup>2</sup>Institut Arthur Vernes, Service de chirurgie cervico-faciale et ORL, Paris, France

<sup>3</sup>Service de Chirurgie Digestive, Bariatrique et Endocrinienne, Hôpitaux Universitaires Paris Seine-Saint-Denis, Hôpital Avicenne, Bobigny, France

Après thyroïdectomie totale, la plainte vocale des patients, hors paralysie récurrentielle, peut être attribuée à l'atteinte de la branche crico-thyroïdienne du nerf laryngé supérieur, entraînant une fatigue vocale (Neri et al., 2011), un abaissement du fondamental (Roy et al., 2009), une difficulté d'atteinte des aigus (Potenza et al., 2017), une voix affaiblie, soufflée et monotone (Roy et al., 2009), et des difficultés de modulation vocale (Orestes & Chhetri, 2014) impactant la prosodie de questions totales, moins montantes (Le Pape et al., 2021). Des mesures de fréquence fondamentale ( $f_0$ ) et de CPPS (Maryn et al., 2009) sont-elles modifiées en fonction de la tâche assignée aux patients (lecture vs répétition d'un même texte) ?

30 femmes francophones avec suspicion d'atteinte de la branche crico-thyroïdienne du NLS (CT- : âge moyen : 56,2 ans, 26- 74 ans), comparées à 30 autres sans dommage (CT+ : âge moyen 57,2 ans, 24-86 ans) après thyroïdectomie totale, au regard d'une population témoin de 29 femmes (T : âge moyen : 55,9 ans, 26-82 ans), ont été enregistrées sur la lecture et la répétition d'une voix masculine produisant un extrait de texte de 52 mots de l'histoire pour enfants *La Poupee Rouge* (Gripari, 1994). La lecture et la répétition de ce texte dialogué ont permis la mesure sous Praat des  $f_0$  moyenne ( $f_0$ moy) et médiane ( $f_0$ med), de son écart-type (ET $f_0$ ), des  $f_0$  minimale ( $f_0$ min) et maximale ( $f_0$ max), de la plage de variation de  $f_0$ , ainsi que du CPPS. Les valeurs de  $f_0$  ont été vérifiées manuellement et la différence entre la lecture et la répétition a été calculée pour chaque sujet et chaque mesure.

Les  $f_0$ moy,  $f_0$ med et ET $f_0$  sont significativement moins différentes en répétition par rapport à la lecture pour les patientes CT- par rapport aux patientes CT+ et aux témoins (*Figure 1*). Pour les patientes CT- et surtout les témoins, les  $f_0$ max,  $f_0$ moy et  $f_0$ med, et la plage de variation de  $f_0$  sont moins élevées en répétition du modèle masculin.  $f_0$ max et la plage de variation de  $f_0$  sont très variables. Le CPPS ne montre aucune tendance.

Ces résultats montrent une moins grande modification de plusieurs paramètres mesurés chez les patientes CT- en répétition par rapport à la lecture (différence lecture-

répétition moins importante). Ils seront discutés en termes de flexibilité pour compenser un défaut de modulation vocale. Le travail orthophonique au niveau de cette compensation pourra être envisagé à la lumière de ces premiers constats.

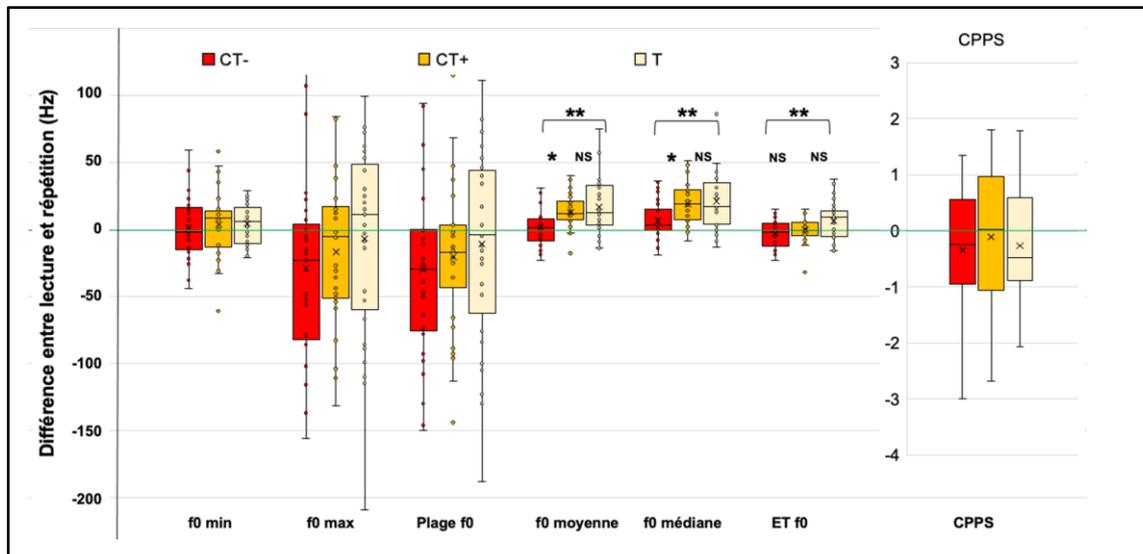


Figure 1. Moyenne (x), médiane (trait horizontal) et variabilité de la différence entre la lecture et la répétition de : f0 minimale, f0 maximale, la plage de variation de f0, f0 moyenne, f0 médiane, l'écart-type de F0 et le CPPS des patientes post-thyroïdectomie avec suspicion d'atteinte du NLS (CT-), sans cette suspicion (CT+) et des femmes témoin (T).

## Références bibliographiques

- Gripari, A. (1994). *Nouveaux contes* [CD audio]. Paris : Compacts Radio France.
- Le Pape, G., Lazard, D. S., Gatignol, P., Tresallet, C., & Pillot-Loiseau, C. (2021). Voice modulation, self-perception and motor branch of the superior laryngeal nerve. *European annals of otorhinolaryngology, head and neck diseases*, 138(4), 241-245. <https://doi.org/10.1016/j.anorl.2020.11.007>
- Neri, G., Castiello, F., Vitullo, F., De Rosa, M., Ciammetti, G., & Croce, A. (2011). Post-thyroidectomy dysphonia in patients with bilateral resection of the superior laryngeal nerve: a comparative spectrographic study. *Acta Otorhinolaryngologica Italica*, 31(4), 228-234.
- Orestes, M. I., & Chhetri, D. K. (2014). Superior laryngeal nerve injury: effects, clinical findings, prognosis, and management options. *Current opinion in otolaryngology & head and neck surgery*, 22(6), 439-443. doi: 10.1097/MOO.0000000000000097
- Maryn, Y., Roy, N., De Bodt, M., Van Cauwenberge, P., & Corthals, P. (2009). Acoustic measurement of overall voice quality: A meta-analysis. *The Journal of the Acoustical Society of America*, 126(5), 2619-2634. <https://doi.org/10.1121/1.3224706>

- Potenza, A. S., Araujo Filho, V. J., & Cernea, C. R. (2017). Injury of the external branch of the superior laryngeal nerve in thyroid surgery. *Gland surgery*, 6(5), 552-562. doi: 10.21037/gs.2017.06.15
- Roy, N., Barton, M. E., Smith, M. E., Dromey, C., Merrill, R. M., & Sauder, C. (2009). An in vivo model of external superior laryngeal nerve paralysis: laryngoscopic findings. *The Laryngoscope*, 119(5), 1017-1032. <https://doi.org/10.1002/lary.20193>

# Exploration des stratégies compensatoires respiratoires en parole : le cas des fentes palatines

Delphine CHARUAU<sup>1,2</sup>  
Béatrice VAXELAIRE<sup>2</sup>  
Rudolph SOCK<sup>2,3</sup>

<sup>1</sup>GIPSA-lab UMR 5216 CNRS, Grenoble INP  
Univ. Grenoble Alpes

<sup>2</sup>Linguistique, Langues et Parole LiLPa,  
Université de Strasbourg

<sup>3</sup>Language, Information and Communication Laboratory  
(LICOLAB),  
Pavol Jozef Šafárik University – Košice

Les fentes palatines sont des malformations congénitales affectant la production de la parole, notamment la réalisation des consonnes (Bechet et al., 2012). Ces troubles articulatoires peuvent être conséquents à une insuffisance vélo-pharyngée entraînant une perte de pression intraorale lors de la phonation (Mayo, Warren & Zajac, 1998). Si ces locuteurs mettent en place des stratégies compensatoires au niveau articulatoire (Harding & Grunwell, 1996), ils mobilisent également de plus grands volumes d'air expiratoires afin de pallier la perte de pression intraorale (Laine et al., 1989 ; Warren, Drake & Davis, 1992). Compte tenu l'interaction étroite entre la respiration et la structuration des énoncés (Winkworth et al., 1995), ces réajustements respiratoires pourraient avoir des conséquences sur l'organisation du discours.

Les objectifs de cette étude sont d'examiner les stratégies compensatoires mises en place au niveau respiratoire par les locuteurs porteurs de fentes palatines, afin de pallier les dysfonctionnements vélo-pharyngés, et d'identifier les effets de ces réajustements sur l'organisation spatio-temporelle du discours dans deux tâches de parole différentes.

Pour cette étude, 17 locuteurs témoins et 12 locuteurs porteurs de fente palatine, âgés de 7 à 12 ans, ont été enregistrés en parole lue et en parole semi-spontanée. Les données respiratoires ont été recueillies à l'aide de ceintures électromagnétiques élastiques (système Respirace ©), disposées sur le thorax et l'abdomen. Synchronisées au signal de parole, enregistré à l'aide d'un microphone, ces données offrent une représentation de la variation des périmètres thoracique et abdominal tout au long de la phonation.

Au vu des résultats, il apparaît que les principaux réajustements déployés par les enfants porteurs de fente palatine sont mis en place lors de l'expiration. En effet, si l'amplitude inspiratoire ne montre aucune différence significative entre les deux groupes de locuteurs, quelle que soit la tâche de parole, la compression de l'ensemble thoraco-abdominal tend à être plus élevée chez les locuteurs porteurs de fente palatine, en comparaison au groupe contrôle (*Figure 1*). En outre, les enfants porteurs de fente terminent leurs groupes de souffle en puisant davantage dans leur volume expiratoire de réserve, en comparaison avec les locuteurs témoins, notamment en parole semi-spontanée.

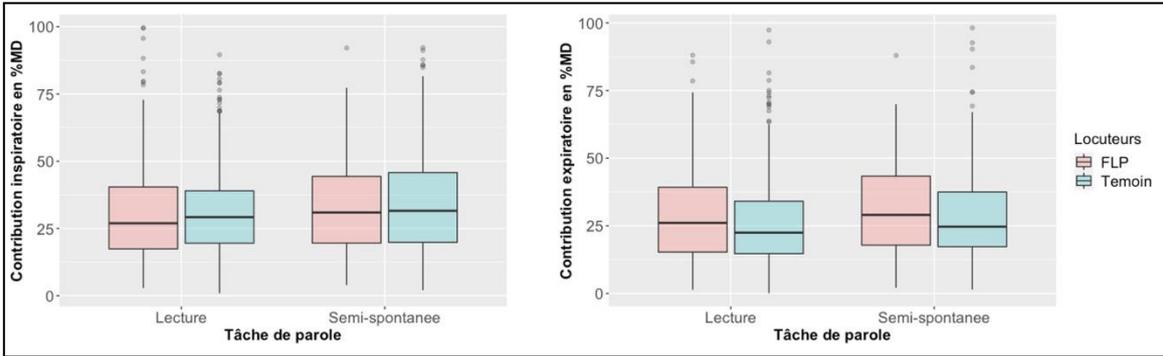


Figure 1. Valeurs de l'amplitude inspiratoire (à gauche) et de l'amplitude expiratoire (à droite) selon le groupe de locuteurs et selon la tâche de parole

En lecture, la durée des cycles respiratoires tend à être plus courte chez les locuteurs porteurs de fente palatine, notamment en raison d'une réduction des groupes de souffle, tandis qu'en parole semi-spontanée, leur durée est similaire entre les deux groupes de locuteurs (Figure 2). Le nombre de syllabes par groupe de souffle est également réduit chez les enfants avec fente. La différence mesurée entre ces deux tâches de parole sera discutée lors du colloque.

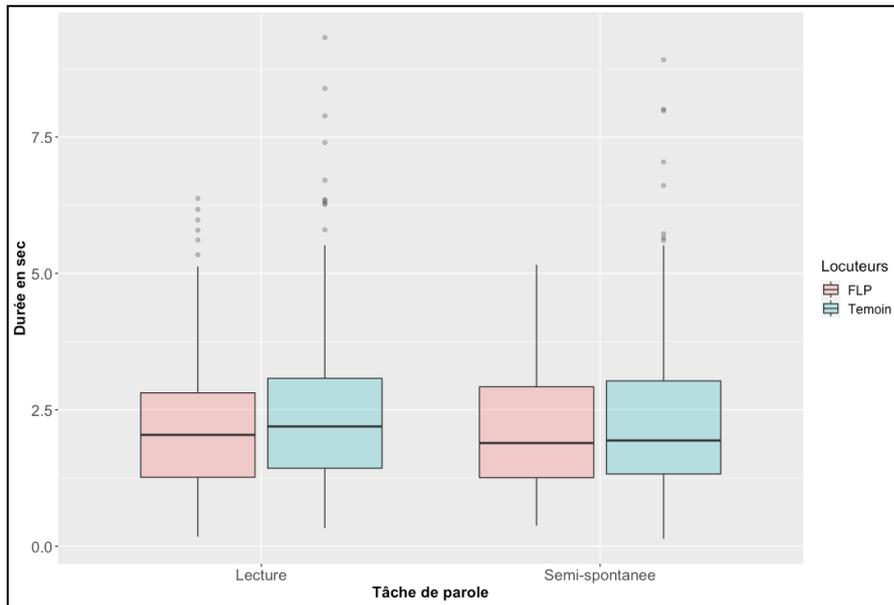


Figure 2. Durée des groupes de souffle (sec) chez les locuteurs témoins et porteurs de fente palatine, en parole lue et semi-spontanée

## Références bibliographiques

- Bechet, M., Hirsch, F., Fauth, C., & Sock, R. (2012). Consonantal space area in children with a cleft palate: an acoustic study. In INTERSPEECH 2012, Home 13th Annual Conference of the International Speech Communication Association, Portland, 58-61.
- Harding, A., & Grunwell, P. (1996). Characteristics of cleft palate speech. *International Journal of Language and Communication Disorders*, 31(4), 331-357.
- Laine, T., Warren, D. W., Dalston, R. M., & Morr, K. E. (1989). Effects of velar resistance on speech aerodynamics. *European Journal of Orthodontics*, 11(7), 52-58.
- Mayo, R., Warren, D. W., & Zajac, D. J. (1998). Intraoral Pressure and Velopharyngeal Function. *The Cleft Palate-Craniofacial Journal*, 35(4), 299- 303.
- Warren, D. W., Drake, A. F., & Davis J. U. (1992). Nasal airway in breathing and speech. *The Cleft Palate-Craniofacial Journal*, 29(6), 511-519.
- Winkworth, A. L., Davis, P. J., Adams, R. D., & Ellis E. (1995). Breathing patterns during Spontaneous Speech. *Journal of Speech, Language, and Hearing Research*, 38(1), 124-144.

# Contribution de MonPaGe-2.0.s au diagnostic différentiel de troubles moteurs acquis de la parole

Cécile FOUGERON<sup>1</sup>  
Frédéric ASSAL<sup>2</sup>  
Angéline BOURBON<sup>1</sup>  
Ina KODRASTI<sup>3</sup>  
Nathalie LEVEQUE<sup>1,4</sup>  
Marina LAGANARO<sup>5</sup>  
Michaela PERNON<sup>1,6</sup>

<sup>1</sup>Laboratoire de Phonétique et Phonologie, UMR7018, CNRS – Université Sorbonne Nouvelle, Paris, France

<sup>2</sup>Hôpital Universitaire de Genève, Suisse

<sup>3</sup>Signal Processing for Communication Group, Idiap Research Institute, Martigny, Suisse

<sup>4</sup>Centre Référent SLA, Département de Neurologie, Hôpital Pitié-Salpêtrière - APHP, Paris, France

<sup>5</sup>Laboratoire de Psycholinguistique, Faculté de Psychologie et de Sciences de l'Éducation, Université de Genève, Suisse

<sup>6</sup>CRMR Wilson & Unité Parkinson, Service de Neurologie, Hôpital Fondation A. de Rothschild, Paris, France

L'outil informatisé d'évaluation clinique de la parole MonPaGe-2.0.s vise à diagnostiquer des troubles moteurs acquis de la parole (MSDs) chez des locuteurs francophones. Sur la base d'une combinaison de scores acoustiques et perceptifs exprimés en 'scores de déviance' par rapport à des valeurs de référence établies sur plus de 400 locuteurs, nous avons pu montrer que l'outil est suffisamment sensible pour diagnostiquer des dysarthries légères à modérées par rapport à des productions de locuteurs neurotypiques (Fougeron et al., 2018 ; Laganaro et al., 2021 ; Pernon et al., 2020).

Très récemment (Fougeron, Kodrasi & Laganaro, 2022), nous avons commencé à questionner l'intérêt de l'outil pour le diagnostic différentiel, i.e. sa capacité à distinguer différentes MSDs, sur la base de ces mêmes scores. Ces derniers capturent des déviations en termes d'intelligibilité, d'articulation, de voix, de débit de parole, de temps maximum de phonation, de contraste prosodique et de débit diadococinésique. L'étude présentée ici s'inscrit en continuation de ce travail.

Dans un premier temps, 60 locuteurs présentant des MSDs légers à modérés ont été répartis en 6 groupes de 10 en rapport avec leur pathologie neurologique (sclérose latérale amyotrophique (SLA), maladies de Wilson, de Parkinson, de Kennedy, ataxie spinocérébelleuse (SCA), apraxie post-AVC). Les résultats de la classification automatique

par paires montrent que la discrimination entre les groupes s'appuie principalement sur les scores d'articulation, de débit (parole et DDK), et voix. Certains groupes sont mieux discriminés que d'autres, avec plus de 85% de classification correcte pour les groupes SCA, maladie de Parkinson et maladie de Kennedy. La classification des locuteurs des autres groupes est légèrement inférieure (79% à 82%), ce qui pourrait en partie s'expliquer par l'association de plusieurs MSDs dans ces catégories cliniques non pures.

### Références bibliographiques

- Fourgeron, C., Delvaux, V., Menard, L., & Laganaro, M. (2018). The MonPaGe\_HA database for the documentation of spoken French throughout adulthood. In: *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*.
- Fougeron, C., Kodrasi, I., & Laganaro, M. (2022). Differentiation of Motor Speech Disorders through the Seven Deviance Scores from MonPaGe-2.0. *Brain Sciences*, 12(11), 1471. <https://doi.org/10.3390/brainsci12111471>
- Laganaro, M., Fougeron, C., Pernon, M., Leveque, N., Borel, S., Fournet, M., Catalano Chiuve, S., Lopez, U. Trouville, R., Menard, L. (2021). Sensitivity and specificity of an acoustic-and perceptual-based tool for assessing motor speech disorders in French: The MonPaGe-screening protocol. *Clinical Linguistics & Phonetics*, 35(11), 1060-1075. <https://doi.org/10.1080/02699206.2020.1865460>
- Pernon, M., Leveque, N., Delvaux, V., Assal, F., Borel, S., Fougeron, C., Trouville, R., Laganaro, M. (2020). MonPaGe, un outil de screening francophone informatise d'évaluation perceptive et acoustique des troubles moteurs de la parole (dysarthries, apraxie de la parole). *Rééducation orthophonique*, 281, 169-197.

# Vers une interprétation, en terme d'altération de traits phonétiques, du niveau d'intelligibilité chez des patients atteints de cancer de la tête ou du cou

Sondes ABDERRAZEK<sup>1</sup>  
Corinne FREDOUILLE<sup>1</sup>  
Alain GHIO<sup>2</sup>  
Muriel LALAIN<sup>2</sup>  
Christine MEUNIER<sup>2</sup>  
Virginie WOISARD<sup>3</sup>

<sup>1</sup>LIA, Avignon Université

<sup>2</sup>Aix-Marseille Univ, LPL, CNRS, Aix-en-Provence

<sup>3</sup>UT2J, LNPL, Toulouse Université & Toulouse Hospital

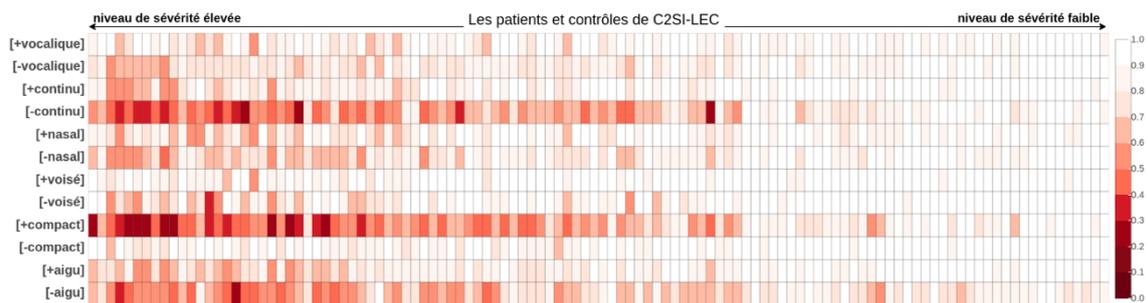
Récemment, nous avons proposé un cadre analytique général, appelé Neuro-based Concept Detector (NCD), pour interpréter les représentations profondes d'un réseau de neurones (DNN). Basé sur les schémas d'activation des neurones cachés, ce cadre met en évidence la capacité des neurones à détecter un concept spécifique lié à la tâche finale de classification dévolue au DNN. Son principal atout est de fournir un outil d'interprétabilité générique pour tout type de DNN quel que soit le domaine d'application. Dans le domaine qui nous intéresse ici – la phonétique clinique, nous avons démontré l'émergence de traits phonétiques (« concept » évoqué plus haut) dans les couches de classification d'un modèle basé sur une architecture de type CNN (Convolutional Neural Network) pour une tâche de classification des phonèmes du français sur de la parole lue et saine. Appliquée à de la parole dégradée produite par des patients atteints d'un cancer de la tête ou du cou, nous avons montré que cette structure reflète automatiquement le niveau d'altération des traits phonétiques des productions des patients.

Le travail décrit ci-dessus fait partie d'un projet à long terme<sup>1</sup> qui vise à déterminer les unités linguistiques qui contribuent le plus au maintien ou à la perte d'intelligibilité dans les troubles de la parole (Woisard et al., 2021). Trois étapes ont été identifiées pour atteindre cet objectif : (Step 1) Modélisation des caractéristiques des unités phonémiques de la parole "normale" grâce à un CNN dédié à une tâche de classification des phonèmes du français (Abderrazek et al., 2020), (Step 2) étude des propriétés de représentation du modèle profond en termes de contenu phonétique (Abderrazek et al., 2022a, Abderrazek et al., 2022b), (Step 3) transfert de cette modélisation et de ses propriétés de représentation dans une tâche de prédiction de l'intelligibilité, typiquement dans le contexte de parole normale et dégradée, en

---

<sup>1</sup> Ce travail a été réalisé dans le cadre du projet RUGBI ("Looking for Relevant linguistic Units to improve the intelligibility measurement of speech production disorders") financé par l'Agence Nationale de la Recherche (Grant N° ANR-18CE45-0008-04).

vue d'étudier sa capacité à fournir une interprétation de la contribution des unités phonémiques dans l'intelligibilité et sa variation (amélioration ou altération). Il s'agit ici de présenter nos avancées sur la réalisation de l'étape 3 à savoir l'intégration de l'approche NCD décrite ci-dessus dans un système automatique de prédiction de l'intelligibilité. Les premiers résultats montrent une corrélation autour de 0.85 entre les scores d'intelligibilité prédits automatiquement par le système développé en étape 3 et ceux évalués perceptivement par les experts. L'utilisation conjointe des approches développées dans les étapes 1 et 2 avec le système de prédiction de l'étape 3 permet de mettre en relation un score d'intelligibilité prédit avec un niveau d'altération des traits phonétiques du français, issu de l'approche NCD, pour interprétation par des experts cliniciens (voir *Figure 1*).



*Figure 1.* Variation des scores issus de l'étape 2 représentant le niveau d'altération mesuré automatiquement par trait phonétique des consonnes (ordonnée) et par locuteur - patients et sujets contrôle (abscisse) triés par mesure de sévérité - à gauche niveau le plus sévère.

Exemple de représentation à mettre en relation avec les scores d'intelligibilité prédits automatiquement par le système (étape 3).

## Références bibliographiques

- Abderrazek, S., Fredouille, C., Ghio, A., Lalain, M., Meunier, C., & Woisard, V. (2020). *Towards Interpreting Deep Learning Models to Understand Loss of Speech Intelligibility in Speech Disorders Step 1: CNN model-based phone classification*. Interspeech'20, Shanghai, China.
- Abderrazek, S., Fredouille, C., Ghio, A., Lalain, M., Meunier, C., & Woisard, V. (2022a). *Towards interpreting deep learning models to understand loss of speech intelligibility in speech disorders Step 2: contribution of the emergence of phonetic traits*. ICASSP, Singapore.
- Abderrazek, S., Fredouille, C., Ghio, A., Lalain, M., Meunier, C., & Woisard, V. (2022b). *Validation of the Neuro-Concept Detector framework for the characterization of speech disorders: A comparative study including Dysarthria and Dysphonia*. Interspeech'22, Corée.
- Woisard, V., Astésano, C., Balaguer, M., Farinas, J., Fredouille, C., Gaillard, P., Ghio, A., Giusti, L., Laaridh, I., Lalain, M., Lepage, B., Mauclair, J., Nocaudie, O., Pinquier, J., Pouchoulin, G., Puech, M., Robert, D., & Roger, V. (2021b). C2SI corpus : a database of speech disorder productions to assess intelligibility and quality of life in head and neck cancers. *Language Resources and Evaluation*, 55(1), 173-190. <https://doi.org/10.1007/s10579-020-09496-3>

# Une approche granulaire pour la prédiction automatique de l'intelligibilité de la parole

Sebastião QUINTAS<sup>1</sup>  
Julie MAUCLAIR<sup>1</sup>  
Julien PINQUIER<sup>1</sup>

<sup>1</sup>IRIT, Université de Toulouse, CNRS, Toulouse INP, UT3, Toulouse, France

La prédiction automatique de l'intelligibilité de la parole peut être considérée comme une alternative pertinente aux évaluations perceptives réalisées en clinique. Ces évaluations sont connues pour être subjectives, biaisées et variables, puisque l'évaluation peut être conditionnée par une variété d'aspects, tels que la connaissance préalable du patient ou la tâche d'évaluation de la parole elle-même (Balaguer et al., 2019). Dans ces conditions, le développement d'une approche automatique peut être considéré comme une alternative plus robuste, reproductible et objective.

Cependant, les évaluations automatiques peuvent être difficiles à interpréter, un aspect qui n'aide pas à l'adaptation clinique de celles-ci. Ainsi, dans cette étude, nous présentons une approche automatique de prédiction de l'intelligibilité, en utilisant les niveaux de granularité de la phrase, du mot et du phonème. Au niveau des phrases, le paradigme des *X-vector speaker embeddings* est utilisé (Quintas et al., 2020). Au niveau des mots, un mécanisme d'auto-attention est appliqué à la tâche de lecture des pseudo-mots du même corpus (Quintas et al., 2022a). Finalement, au niveau des phonèmes, un réseau siamois est appliqué pour prédire un score d'intelligibilité en utilisant la similarité phonémique (Quintas et al., 2022b). La fusion de ces trois systèmes, voir *Figure 1*, crée une couche supplémentaire d'interprétation de grande pertinence clinique. Ce travail est appliqué au corpus français du cancer des voies aérodigestives supérieures (Woisard et al., 2020). Les scores d'intelligibilité de référence, obtenus dans le cadre de ce corpus, correspondent à l'évaluation perceptive de six juges professionnels. Ils sont compris entre 0 (inintelligible) et 10 (parfaitement intelligible).

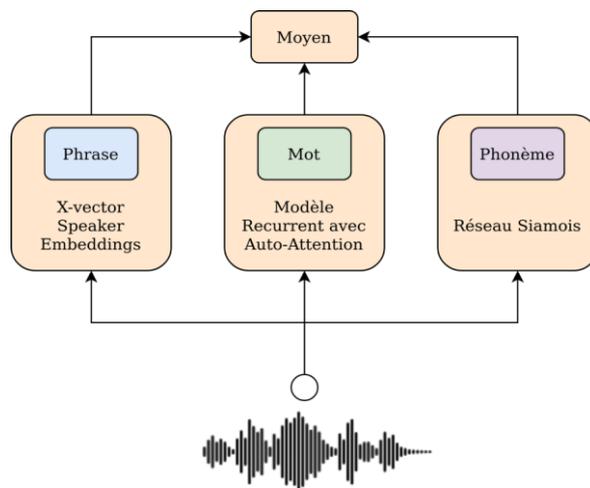


Figure 1. Schéma de l'approche granulaire proposée et de la fusion

Les résultats, illustrés dans le *Tableau 1*, montrent la corrélation et l'erreur (RMSE) obtenue avec chaque système ainsi que leur combinaison. Individuellement, il est possible d'obtenir de hautes corrélations pour chaque système. Toutefois, avec leur fusion, le système est encore plus performant. La comparaison avec un système « oracle » (qui choisit manuellement le meilleur système pour chaque patient) montre que des améliorations sont encore possibles, notamment au niveau de l'erreur. Cette différence est principalement due à la sous-performance de notre système sur les patients à faible intelligibilité, une classe sous-représentée dans le corpus.

L'approche granulaire permet l'analyse de plusieurs niveaux qui sont intéressants pour les cliniciens et permet l'interprétabilité par des humains du fonctionnement du système automatique. Si le système automatique fournit des scores interprétables, cela favorisera son acceptabilité par les thérapeutes.

*Tableau 1.* Corrélations (Spearman) et erreur quadratique moyenne (RMSE) obtenues avec les systèmes individuels ainsi que leur fusion

Granularité	Système		RMSE
Phrase	X-Vector Speaker Embeddings (Quintas, S. 2020)	0,85	1,623
Mot	Modèle récurrent avec auto-attention (Quintas, S. 2022a)	0,82	1,548
Phonème	Réseau Siamois (Quintas, S. 2022b)	0,89	2,080
<b>Fusion des 3 systèmes (score moyen)</b>		<b>0,91</b>	<b>1,588</b>
<i>Oracle (Choix manuel du meilleur score)</i>		<i>0,92</i>	<i>1,075</i>

### Références bibliographiques

- Balaguer, M., Pommée, T., Farinas, J., Pinquier, J., Woisard, V., & Speyer, R. (2019). Effects of oral and oropharyngeal cancer on speech intelligibility using acoustic analysis : Systematic review. *Head & neck*, 42(1), 111-130. <https://doi.org/10.1002/hed.25949>
- Woisard, V., Astésano, C., Balaguer, M., Farinas, J., Fredouille, C., Gaillard, P., Ghio, A., Giusti, L., Laaridh, I., Lalain, M., Lepage, B., Mauclair, J., Nocaudie, O., Pinquier, J., Pouchoulin, G., Puech, M., Robert, D., & Roger, V. (2021). C2SI corpus : a database of speech disorder productions to assess intelligibility and quality of life in head and neck cancers. *Language Resources and Evaluation*, 55(1), 173-190. <https://doi.org/10.1007/s10579-020-09496-3>
- Quintas, S., Abad, A., Mauclair, J., Woisard, V., Pinquier, J. (2022a). Utilisation de modèles transformers pour la prédiction de l'intelligibilité de la parole de patients atteints de cancers des voies aérodigestives supérieures. Journées d'Études sur la parole.
- Quintas, S., Mauclair, J., Woisard, V., & Pinquier, J. (2022b). Automatic Assessment of Speech Intelligibility using Consonant Similarity for Head and Neck Cancer. Proceedings of Interspeech.
- Quintas, S., Mauclair, J., Woisard, V., & Pinquier, J. (2020). Automatic Prediction of Speech Intelligibility Based on X-Vectors in the Context of Head and Neck Cancer. Proceedings of Interspeech.

# Posters



# Session posters 1



# Utilisation de l'échographie comme outil d'évaluation des troubles des sons de la parole (TSP) dans un contexte clinique québécois : étude pilote sur la connaissance de l'outil et la variabilité des analyses

Isabelle DÉMOSTHÈNES<sup>1</sup>  
Lucie MÉNARD<sup>1</sup>

<sup>1</sup>Laboratoire de Phonétique,  
Université du Québec à Montréal

Selon Cleland et coll. (2020) et Sugden et Cleland (2021), l'ajout de l'échographie lors de la transcription phonétique diminuerait la variabilité interjuge comparativement à une transcription basée sur le signal audio uniquement. Toutefois, très peu de données quant à la connaissance de l'outil et l'interprétation des images dans le contexte clinique québécois sont disponibles.

Nous avons donc sollicité, 39 orthophonistes, provenant de 11 régions administratives et de divers milieux, pour compléter un sondage en ligne (tâche 1). Les questions visaient à déterminer si les répondants avaient entendu parler de l'utilisation de l'échographie en orthophonie ainsi que leur niveau d'exposition aux images échographiques dans leur pratique.

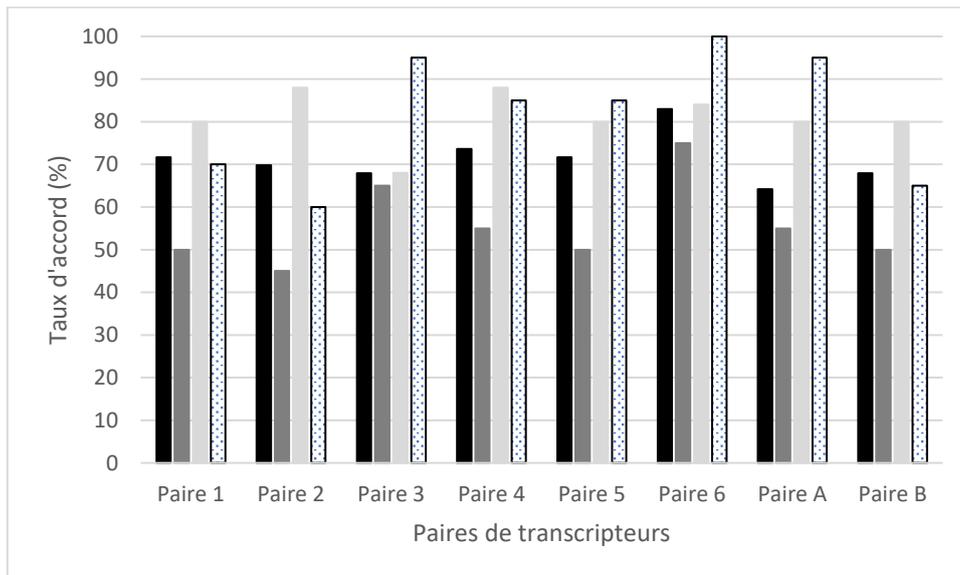
Par la suite, 13 cliniciens ont été recrutés pour transcrire et coter 40 vidéos de mots isolés (tâche 2). Ces vidéos (20 sur le plan sagittal et 20 sur le plan coronal) avaient été recueillis en clinique, auprès de deux enfants ayant un TSP, à l'aide d'un échographe Sonosite Edge II couplé à une sonde P19, 5,1 MHz (Fujifilm SonoSite Inc., 2017). Une formation préalable de 3 heures sur l'échographie et les méthodes d'analyse de Cleland et coll. (2018) a été donnée aux participants via *Zoom*.

Dans un troisième temps, 5 cliniciens supplémentaires ont réalisé une transcription et une cotation à l'aide uniquement du signal audio extrait des vidéos utilisés précédemment (tâche 3).

Nous avons exploré la variabilité des analyses dans les tâches 2 et 3 en mesurant l'accord interjuge sur 20 mots de quatre façons :

- Taux d'accord point par point sur la transcription phonétique entre deux participants appariés au sein de la même modalité ;
- Dans cette transcription, taux d'accord sur le trait ciblé par l'échographie pour les 20 consonnes dont l'image était à coter (consonnes sélectionnées en lien avec les difficultés articulatoires de chaque cas) ;
- Kappa de Fleiss sur la transcription en API des 53 phonèmes ;
- Kappa de Fleiss sur l'appréciation de la réalisation articulatoire (adéquate ou non) des 20 consonnes cotées par les participants.

Avec un taux d'accord moyen sur 53 phonèmes de 72,95 % (écart-type, 5,29 %) pour les six paires ayant accès à l'échographie et de 66,05 % (écart-type, 2,62 %) pour les deux utilisant seulement le signal audio, nos résultats ne nous permettent pas de confirmer que l'ajout de l'échographie améliore la transcription phonétique des cliniciens (voir *Figure 1* et *Tableau 1*). Les différences dans les devis expérimentaux et l'expérience dans l'interprétation des images échographiques pourraient expliquer ces résultats.



*Figure 1.* Taux d'accord interjuge dans la transcription phonétique pour chaque paire de transcripteurs. Paires 1 à 6, transcription avec le support visuel des vidéos de l'échographie (tâche 2). Paires A et B, transcription avec le signal audio seulement (tâche 3).

- Taux pour l'ensemble des 53 phonèmes (20 mots) présentés pour l'analyse interjuge
- Taux pour les 20 consonnes ciblées pour l'analyse de la production
- Taux pour les 25 voyelles présentées
- ▨ Taux pour le trait augmenté par l'image échographique pour les 20 consonnes ciblées

Tableau 1. Niveau d'accord interjuge pour l'ensemble des participants

	Transcripteurs ayant accès au support visuel de l'échographie (n = 13)	Transcripteurs ayant accès au signal audio seulement (n = 5)
Kappa de Fleiss sur la transcription en API des 53 phonèmes	Bonne concordance kappa = 0,721 (IC 95 %, 0,712 à 0,730), p. < 0,001	Concordance modérée à bonne kappa = 0,605 (IC 95 %, 0,584 à 0,627), p. < 0,001
Kappa de Fleiss sur la réalisation articulaire (adéquate ou non) des 20 consonnes ciblées	Légère concordance kappa = 0,332 (IC 95 %, 0,278 à 0,386), p. < 0,001	Légère concordance kappa = 0,329 (IC 95 %, 0,187 à 0,471), p. < 0,001

### Références bibliographiques

- Cleland, J., Lloyd, S., Campbell, L., Crampin, L., Palo, J. P., Sugden, E., Wrench, A., & Zharkova, N. (2020). The Impact of Real-Time Articulatory Information on Phonetic Transcription: Ultrasound-Aided Transcription in Cleft Lip and Palate Speech. *Folia Phoniatrica et Logopaedica*, 72(2), 120-130. <https://doi.org/10.1159/000499753>
- Cleland, J., Wrench, A., Lloyd, S., & Sugden, E. (2018). ULTRAX2020 : Ultrasound Technology for Optimising the Treatment of Speech Disorders : Clinicians' Resource Manual (Issue February). <https://doi.org/10.15129/63372>
- Shriberg, L. D., & Lof, G. L. (1991). Reliability studies in broad and narrow phonetic transcription. *Clinical Linguistics & Phonetics*, 5(3), 225-279. <https://doi.org/10.3109/02699209108986113>
- Sugden, E., & Cleland, J. (2021). Using ultrasound tongue imaging to support the phonetic transcription of childhood speech sound disorders. *Clinical Linguistics & Phonetics*, 1-20. <https://doi.org/10.1080/02699206.2021.1966101>
- Viera, A. J., & Garrett, J. M. (2005). Understanding interobserver agreement: The Kappa statistic. *Family Medicine*, 37(5), 360-363.



# Relations entre évaluation acoustique et perceptive de la qualité de la voix : une approche dynamique

Nicolas AUDIBERT<sup>1</sup>  
Claire PILLOT-LOISEAU<sup>1</sup>  
Lise CREVIER-BUCHMAN<sup>1,3</sup>  
Sophie FAGNIART<sup>2</sup>  
Bernard HARMEGNIES<sup>2</sup>  
Kathy HUET<sup>2</sup>  
Amelia PETTIROSSI<sup>1,4</sup>  
Myriam PICCALUGA<sup>2</sup>  
Reina REMMAN<sup>1</sup>  
Virginie ROLAND<sup>2</sup>  
Lola TERNY<sup>2</sup>  
Clémence VERHAEGEN<sup>2</sup>  
Véronique DELVAUX<sup>2</sup>

<sup>1</sup>Laboratoire de Phonétique et phonologie UMR 7018, CNRS/Université Sorbonne  
Nouvelle, Paris, France.

<sup>2</sup>Service de Métrologie et Sciences du Langage,  
Institut de Recherche en Sciences et Technologies du Langage & Fond National de La  
Recherche Scientifique,  
Université de Mons, Belgique

<sup>3</sup>Hôpital Foch : Service de Laryngologie Phoniatre,  
92150 Suresnes, France

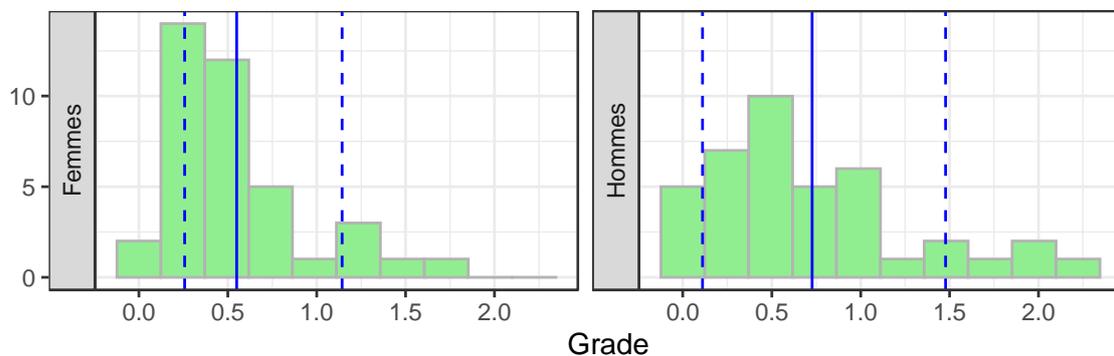
<sup>4</sup>Imagerie & Cerveau,  
UMR1253 INSERM/Université de Tours,  
37000 Tours, France

La notion de qualité vocale (QV) est complexe et multiparamétrique. En contexte clinique, on recommande typiquement une évaluation fondée sur une diversité d'outils/indicateurs (Kreiman & Gerratt, 2010 ; Barsties & De Bodt, 2015 ; Brockmann-Bauser & De Paula Soares, 2023), entre autres : évaluation perceptive par le clinicien à l'aide d'échelles (GRBASI, CAPE-V), auto-évaluation par le patient (VHI), indicateurs acoustiques simples (jitter, shimmer, HNR, CPPs, etc.) ou composites (DSI, AVQI), dont l'interprétation est modulée par l'évaluation de la fonction pneumo-phonatoire et l'imagerie des plis vocaux. Deux problématiques non résolues traversent l'histoire de la recherche en évaluation de la QV : (i) la question des relations, non linéaires et non transparentes, entre paramètres acoustiques et jugements perceptifs ; (ii) le lien entre évaluation de la QV et l'échantillon évalué, /a/ tenus vs. échantillons de parole, plus largement la question de la dynamique temporelle de la QV dans la parole.

Dans cette étude exploratoire, nous adressons ces deux questions de façon conjointe en analysant dans quelle mesure un ensemble de paramètres acoustiques dynamiques permettent de prédire les jugements perceptifs sur les mêmes échantillons de parole. Les productions analysées sont issues d'enregistrements effectués par Delvaux et Pillot-Loiseau (2020) à l'aide de MonPaGe (Laganaro *et al.*, 2021), comprenant la phrase « Mélanie vend du lilas » produite par 39 femmes (21-85 ans) et 40 hommes (20-88 ans) normophoniques issus de la région parisienne. Neuf orthophonistes expérimentées ont jugé les phrases à l'aide de l'échelle GRBASI (distribution des grades : *Figure 1*).

Ces jugements perceptifs sont ici exploités et mis en relation avec les analyses acoustiques suivantes : le rapport harmonique sur bruit, le CPPs (Maryn *et al.*, 2009) et le rapport d'énergie entre les bandes 0-1kHz et 1-4kHz, dont la pertinence perceptive a été démontrée (Kitzing, 1986), ont été mesurés sur trois points de chaque segment de la phrase pour 39 femmes et 40 hommes. Des modèles GAM par variable et sexe ont permis de modéliser l'évolution des mesures dans la phrase, en corrigeant l'influence de F0 sur ces mesures.

On constate que le pouvoir prédictif des indices acoustiques par rapport aux altérations perçues évolue au cours du temps (*Figure 2*), c'est-à-dire en fonction du contenu segmental et/ou de la position prosodique, différemment entre hommes vs. femmes (même après correction pour la différence de F0) et de l'indice concerné. Ces résultats seront discutés à la conférence en lien avec la littérature pertinente (par exemple Pouchoulin *et al.*, 2008).



*Figure 1.* Distribution des grades moyens attribués par les 9 orthophonistes expérimentées aux 39 femmes et 40 hommes pris en compte dans l'analyse. Les lignes verticales bleues indiquent la valeur médiane (trait plein) et les quantiles à 10% et 90% (pointillés).

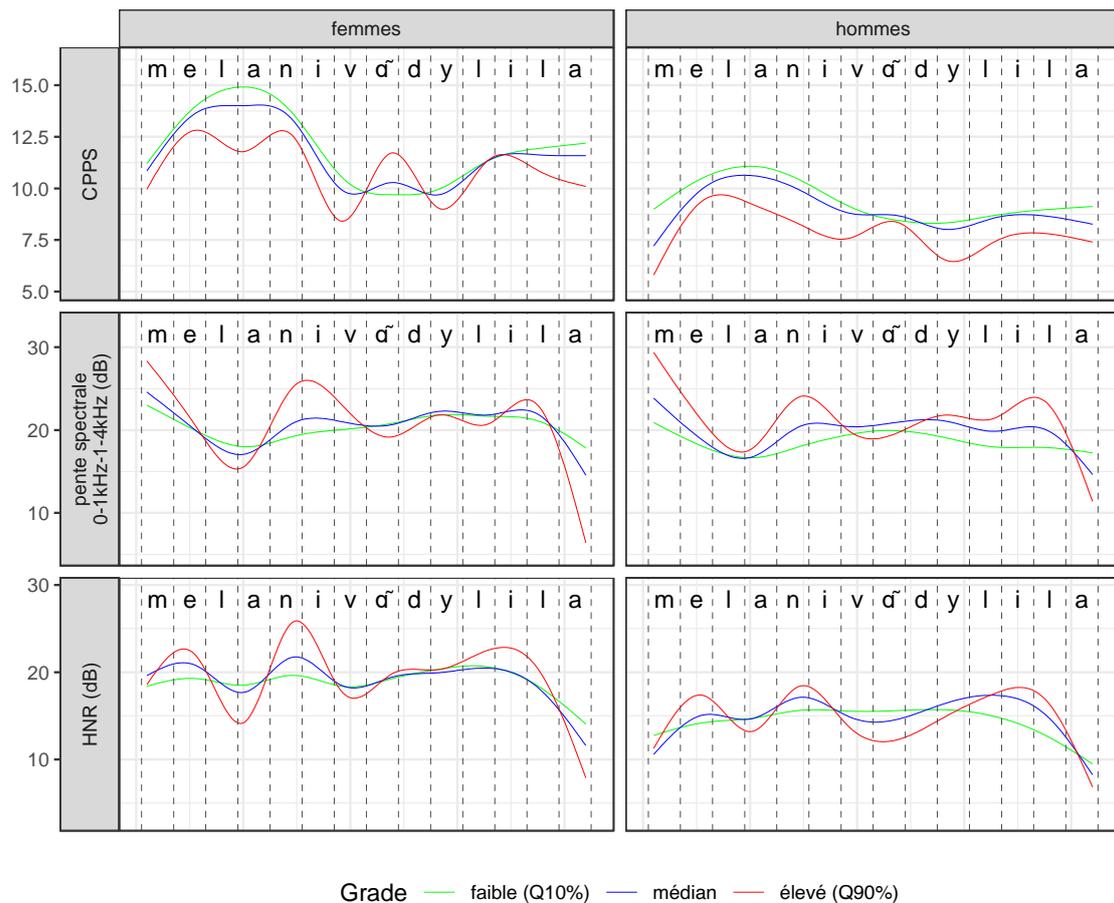


Figure 2. Évolution dans la phrase « Mélanie vend du lilas » de chacune des trois mesures retenues en fonction du grade moyen, obtenue par une modélisation GAM tenant compte des variations de F0, séparément pour hommes et femmes. Les trajectoires sont représentées pour le 1<sup>er</sup> décile (F : 0,14 ; H : 0,11), la médiane (F : 0,43 ; H : 0,56) et le 9<sup>e</sup> décile (F : 1,14, H : 1,44) des valeurs de grade. La couleur de fond rose pâle indique une différence significative uniquement entre grade faible et élevé de cette population sur la portion de signal correspondante, le rose plus foncé une différence significative également entre grade faible et médian.

### Références bibliographiques

- Barsties, B., & De Bodt, M. (2015). Assessment of voice quality : Current state-of-the-art. *Auris Nasus Larynx*, 42(3), 183-188. <https://doi.org/10.1016/j.anl.2014.11.001>
- Brockmann-Bausser, M., & De Paula Soares, M. F. (2023). Do We Get What We Need from Clinical Acoustic Voice Measurements ? *Applied sciences*, 13(2), 941. <https://doi.org/10.3390/app13020941>
- Delvaux, V., & Pillot-Loiseau, C. (2020). Perceptual Judgment of Voice Quality in Nondysphonic French Speakers : Effect of Task-, Speaker- and Listener-Related Variables. *Journal of Voice*, 34(5), 682-693. <https://doi.org/10.1016/j.jvoice.2019.02.013>

- Kitzing, P. (1986). LTAS criteria pertinent to the measurement of voice quality. *Journal of phonetics*, 14(3-4), 477-482.
- Kreiman, J., & Gerratt, B. R. (2010). Perceptual Assessment of Voice Quality : Past, Present, and Future. *Perspectives on voice and voice disorders*, 20(2), 62-67. <https://doi.org/10.1044/vvd20.2.62>
- Laganaro, M., Fougeron, C., Pernon, M., Levêque, N., Borel, S., Fournet, M., Chiuvé, S. C., Lopez, U., Trouville, R., Ménard, L., Burkhard, P. R., Assal, F., & Delvaux, V. (2021). Sensitivity and specificity of an acoustic- and perceptual-based tool for assessing motor speech disorders in French : the MonPaGe-screening protocol. *Clinical Linguistics & Phonetics*, 35(11), 1060-1075. <https://doi.org/10.1080/02699206.2020.1865460>
- Maryn, Y., Roy, N., De Bodt, M., Van Cauwenberge, P., & Corthals, P. (2009). Acoustic measurement of overall voice quality: A meta-analysis. *Journal of the Acoustical Society of America*, 126(5), 2619-2634. <https://doi.org/10.1121/1.3224706>
- Pouchoulin, G., Fredouille C., Bonastre JF, Ghio A, Giovanni A. (2008). Analyse Phonétique dans le Domaine Fréquentiel pour la Classification des Voix Dysphoniques. *Journées d'Étude sur la Parole (JEP 2008)*, Avignon.

### Remerciements

Ce travail a été soutenu par le Partenariat Hubert Curien (PHC) TOURNESOL n°46241WH et par le Laboratoire d'Excellence (LabEx) *Empirical Foundations of Linguistics* (EFL) n°ANR-10-LABX-0083. Il contribue à l'IdEx Université de Paris (ANR-18-IDEX-0001).

# Analyse acoustique de toux volontaires et réflexes : comparaison entre sujets sains et patients avec une tumeur tête et cou

Sofiana MOOTASSIM-BILLAH<sup>1</sup>  
Jean SCHOENTGEN<sup>2</sup>  
Dirk VAN GESTEL<sup>3</sup>

<sup>1</sup> Speech Therapy Unit; Department of Radiation-Oncology ; Institut Jules Bordet, Hôpital  
Universitaire de Bruxelles (HUB), Université Libre de Bruxelles (ULB)

<sup>2</sup> EAMS (Bio-, Electro- And Mechanical Systems); Université Libre de Bruxelles (ULB)

<sup>3</sup> Department of Radiation-Oncology; Institut Jules Bordet, Hôpital Universitaire de  
Bruxelles (HUB),  
Université Libre de Bruxelles (ULB)

La dysphagie radio-induite se définit par des troubles de la déglutition (blocage ou fausse route des aliments vers les voies respiratoires) consécutifs à une radio(chimio)thérapie chez des patients avec une tumeur tête et cou (King et al., 2016). L'objectif de cette étude est d'identifier les indices acoustiques traduisant l'efficacité de la toux volontaire (prédicteur du blocage) et de la toux réflexe (prédicteur de la fausse route) comme mécanismes de protection des voies respiratoires (Troche et al., 2016).

L'analyse temporelle du signal repose sur la durée de la toux, son contour d'amplitude, son contour d'entropie échantillonnée (imprédictibilité) et son contour de kurtosis (impulsivité) (*Figure 1*). Les contours sont rapportés au moyen de trois coefficients : la moyenne, la pente et la courbure. Une deuxième méthode d'analyse rapporte l'énergie par bande de fréquence (0-400Hz, 400-800Hz, 800-1600Hz, 1600-3200Hz, 3200Hz-demi-fréquence d'échantillonnage) (*Figure 2*) ainsi que la prééminence de la périodicité et le jitter pour les bandes 0-400Hz et 400-800Hz (indices de perturbations additives et de modulation).

A l'aide d'un microphone de contact, nous avons récolté par sujet 5 toux volontaires et 4 toux réflexes induites via l'inhalation d'acide citrique auprès de 40 sujets sains et 13 patients présentant des blocages et des fausses routes.

En ce qui concerne les toux volontaires, l'analyse temporelle rapporte que les signaux des patients sont caractérisés par une durée plus longue ( $p < 0.001$ ), une convexité moindre de la courbure d'amplitude ( $p = 0.004$ ), une pente plus abrupte ( $p = 0.009$ ) et une concavité moindre de la courbure ( $p = 0.018$ ) du contour d'entropie échantillonnée ainsi qu'une pente moins abrupte ( $p = 0.037$ ) et une convexité plus marquée de la courbure ( $p = 0.035$ ) du contour du kurtosis. L'analyse spectrale des toux volontaires des patients rapporte une énergie spectrale moindre dans la bande 0-400Hz ( $p = 0.032$ ), une énergie spectrale plus élevée pour les bandes  $> 1600$ Hz ( $p < 0.001$ ), une prééminence de la périodicité moindre  $< 800$ Hz ( $p < 0.001$ ) et des indices de jitter plus élevés  $< 800$ Hz ( $p = 0.009$  ;  $p = 0.045$ ).

En ce qui concerne les toux réflexes, les signaux des patients sont caractérisés par une durée plus courte ( $p=0.005$ ), une courbure du contour d'amplitude de convexité moindre ( $p<0.001$ ) ainsi que par une concavité plus marquée de la courbure ( $p<0.001$ ) du contour d'entropie échantillonnée. L'analyse spectrale des toux réflexes chez les patients met en évidence une énergie spectrale  $>3200\text{Hz}$  plus importante ( $p=0.008$ ), une prépondérance de la périodicité moindre  $<400\text{Hz}$  ( $p<0.001$ ) et un jitter plus élevé  $<400\text{Hz}$  ( $p=0.004$ ).

La conclusion préliminaire est que l'analyse acoustique des toux volontaires et réflexes distingue les sujets sains des patients.

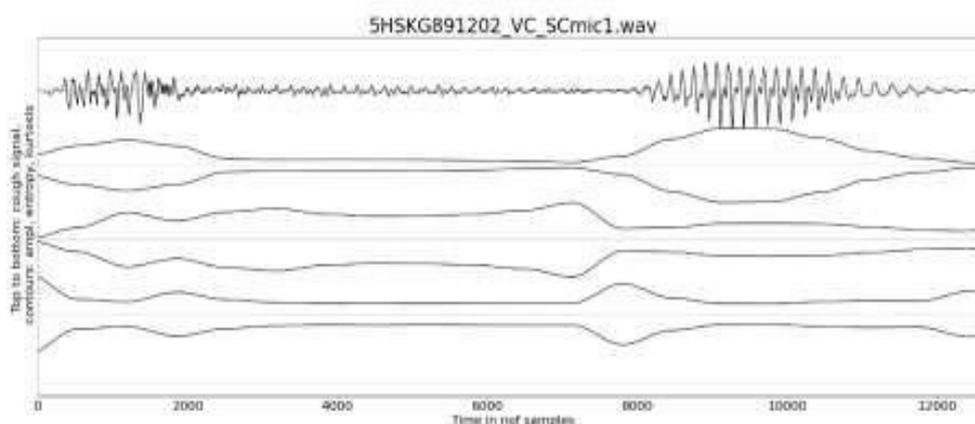


Figure 1. Une émission de toux volontaire d'un sujet sain ; première ligne (en haut) : le signal de toux original ; deuxième ligne : le contour d'amplitude ; troisième ligne : le contour d'entropie échantillonnée ; quatrième ligne : le contour de kurtosis

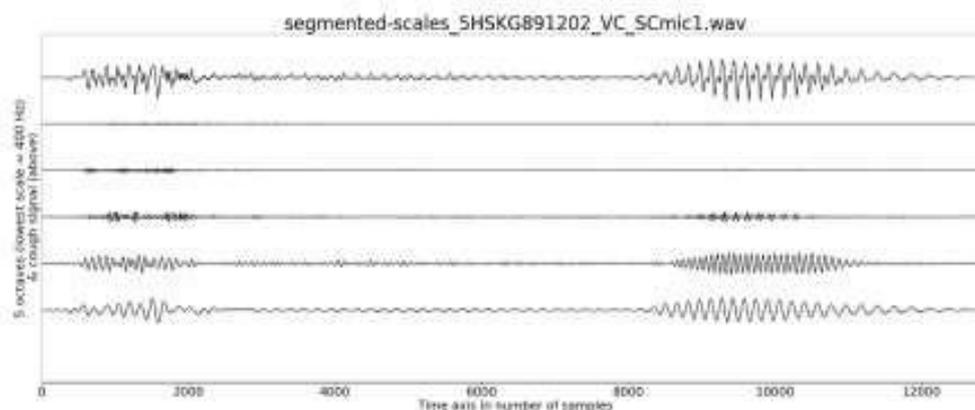


Figure 2. Une émission de toux volontaire d'un sujet sain ; première ligne (en haut) : le signal de toux original ; lignes suivantes : énergie relative rapportée bande par bande (bandes d'octave à partir de 400Hz en bas)

## Références bibliographiques

- King, S. N., Dunlap, N.E., Tennant, P.A. & Pitts, T. (2016). Pathophysiology of Radiation Induced Dysphagia in Head and Neck Cancer. *Dysphagia*, 31(3), 339-351.
- Troche, M. S., Schumann, B., Brandimore, A.E., Okun, M.S, & Hegland, K.W (2016). Reflex Cough and Disease Duration as Predictors of Swallowing Dysfunction in Parkinson's Disease. *Dysphagia*, 31(6), 757–764.



# L'identification des émotions chez des enfants bilingues à partir des indices prosodiques du signal de parole : état de la littérature et proposition méthodologique

Lola TERNY<sup>1</sup>  
Véronique DELVAUX<sup>1,2</sup>  
Virginie ROLAND<sup>1</sup>  
Clémence VERHAEGEN<sup>1</sup>  
Kathy HUET<sup>1</sup>  
Myriam PICCALUGA<sup>1</sup>

<sup>1</sup>Service de Métrologie et Sciences du Langage, Institut de Recherche en Sciences et Technologies du Langage, Université de Mons, Belgique

<sup>2</sup>Fonds National de la Recherche Scientifique, Belgique

Les compétences émotionnelles font référence à la faculté d'identification, de compréhension, d'expression et de gestion de ses propres émotions et de celles d'autrui (Mikolajczak, 2014). Elles connaissent différents stades de développement et un niveau de performance qui s'accroît avec l'âge (Ma et al., 2022 ; Amorim et al., 2021 ; Grobras et al., 2018).

Le traitement des émotions sur les plans perceptif et productif est un phénomène multimodal pour lequel plusieurs canaux sont mobilisés. Si la littérature scientifique est relativement développée et plutôt consensuelle quant au traitement des indices faciaux, elle l'est beaucoup moins en ce qui concerne celui des indices vocaux. Dès lors, ces derniers seraient, dans leur expression, davantage propres à la culture et à la langue de l'individu (Elfenbein & Ambady, 2002) à contrario des expressions faciales plus universelles (Chronaki et al., 2015). De cette façon, les émotions sont mieux identifiées à partir des indices faciaux que vocaux et cette compétence d'identification sur base des indices acoustiques connaît un développement plus tardif au cours de la vie (Aguert et al., 2013 ; Chronaki et al., 2018).

Notre étude porte sur la capacité à identifier les émotions ou l'état émotionnel d'un individu grâce à l'information contenue dans le signal de parole, autrement dit sur base des variations acoustiques produites par l'individu dans un contexte émotionnel (prosodie émotionnelle, PE). L'objectif sera de déterminer, dans le signal de parole, les éléments qui facilitent ou complexifient l'identification d'une émotion parmi d'autres et ce, dès le plus jeune âge, chez les enfants bilingues. En effet, plusieurs recherches ont investigué cette compétence chez une population d'enfants monolingues dans différentes langues (Ma et al., 2022 ; Filipe et al., 2017 ; Morton & Trehub, 2011 ; Aguert et al., 2013) ainsi que chez des adultes apprenants d'une deuxième langue (Zhu 2013 ; Bathara et al., 2016 ; Champoux-Larsson & Dylman, 2021) mais aucune étude à notre connaissance n'a traité ce sujet chez les enfants bilingues.

Nous ferons d'abord état des méthodologies mises en place dans le cadre de précédentes recherches sur la capacité à identifier les émotions à partir de la PE afin de proposer un dispositif original qui évalue cette capacité chez des enfants monolingues francophones. A plus long terme, nous souhaitons poursuivre notre investigation auprès de jeunes enfants bilingues (français et espagnol) afin de mesurer l'influence d'un système prosodique et émotionnel dans une langue sur l'appréhension de celui d'une autre langue.

### Références bibliographiques

- Aguert, M., Laval, V., Lacroix, A., Gil, S., & Le Bigot, L. (2013). Inferring Emotions from Speech Prosody: Not So Easy at Age Five. *PLoS ONE*, 8(12).
- Amorim, M., Anikin, A., Mendes, A. J., Lima, C. F., Kotz, S. A., & Pinhero, A. P. (2021). Changes in vocal emotion recognition across the life span. *Emotion*, 21(2), 315-325.
- Bhatara, A., Laukka, P., Boll-Avetisyan, N., Granjon, L., Anger Elfenbein, H., & Bänziger, T. (2016). Second Language Ability and Emotional Prosody Perception. *PLOS ONE*, 11(6).
- Champoux-Larsson, M.-F., & Dylman, A. S. (2021). Bilinguals' inference of emotions in ambiguous speech. *International Journal of Bilingualism*, 25(5), 1297-1310.
- Chronaki, G., Hadwin, J. A., Garner, M., Maurage, P., & Sonuga-Barke, E. J. S. (2015). The development of emotion recognition from facial expressions and non-linguistic vocalizations during childhood. *British Journal of Developmental Psychology*, 33(2), 218-236.
- Chronaki, G., Wigelsworth, M., Pell, M. D., & Kotz, S. A. (2018). The development of cross-cultural recognition of vocal emotion during childhood and adolescence. *Scientific Reports*, 8(1), 8659.
- Elfenbein, H., & Ambady, N. (2002). Is there an in-group advantage in emotion recognition? *Psychological bulletin*, 128, 243-249.
- Filipe, M. G., Peppé, S., Frota, S., & Vicente, S. G. (2017). Prosodic development in European Portuguese from childhood to adulthood. *Applied Psycholinguistics*, 38(5), 1045-1070.
- Grosbras, M.H., Ross, P. D., & Belin, P., Categorical emotion recognition from voice improves during childhood and adolescence. *Scientific Reports*, 8(1), 14791.
- Ma, W., Zhou, P., & Thompson, W. F. (2022). Children's decoding of emotional prosody in four languages. *Emotion*, 22(1), 198-212.

- Mikolajczak, M. (2014). « Les compétences émotionnelles : historique et conceptualisation »,  
In Mikolajczak, M. *Les compétences émotionnelles*. Dunod, 1-9.
- Morton, J.B. & Trehub, S.E. (2001). Children's understanding of emotion in speech. *Child development*, 72(3), 834-843.
- Zhu, Y. (2013). Which is the best listener group?: Perception of Chinese emotional prosody by Chinese natives, naïve Dutch listeners and Dutch L2 learners of Chinese. *Dutch Journal of Applied Linguistics*, 2(2), 170-183.



# Troubles de la parole dans la maladie de Parkinson : effets d'une prise en charge par l'utilisation de la musique ? Étude de cas

Martin NOPERE<sup>1</sup>  
Véronique DELVAUX<sup>1,2</sup>  
Bernard HARMEGNIES<sup>1</sup>  
Myriam PICCALUGA<sup>1</sup> Kathy HUET<sup>1</sup>  
Virginie ROLAND<sup>1</sup>

<sup>1</sup> Service de Métrologie et Sciences du Langage,  
Institut de Recherche en Sciences et Technologies du Langage, UMONS, Belgique

<sup>2</sup> Fonds National de la Recherche Scientifique, Belgique

Les troubles thymiques et dysarthriques sont des symptômes fréquents dans la maladie de Parkinson (MP). Ces troubles, associés à ceux de la triade parkinsonienne, peuvent entraîner une diminution de l'autonomie et plus largement de la qualité de vie (QoL). Par ailleurs, les traitements médicamenteux les plus fréquemment proposés aux patients peuvent avoir des effets variables sur les troubles dysarthriques, ce qui conduit les professionnels à compléter ceux-ci par des prises en charges (PEC) non-médicamenteuses. La (neuro)musicothérapie (NMT) a ainsi émergé ces dernières années et pourrait apparaître comme une prise en charge innovante et efficace [1, 2]. En effet, des résultats positifs ont été relevés dans les troubles de la marche et de la mobilité chez des patients parkinsoniens (Grand, 2019 ; Koshimori & Thaut, 2018).

Afin d'investiguer si la NMT peut avoir des effets bénéfiques sur la PEC de la dysarthrie hypokinétique (DH), une première étude de cas a été réalisée auprès d'un patient atteint de la MP (76 ans, stade 4 (Goetz, 2010)), qui a été soumis à 6 séances de rééducation par NMT. Ces séances ont consisté à travailler différents aspects de production de la parole (rythme, articulation, intensité). Cette PEC a été précédée et suivie par des évaluations des troubles moteurs de la parole via le protocole de screening MonPaGe [4] et des paramètres thymiques et de la QoL via questionnaires (GAI, GDS15, SHI et PDQ39).

Deux indices acoustiques ont été étudiés : la variabilité de la fréquence fondamentale ( $F_0$ ) (Hz) et le débit de parole (syll./s) via une tâche de description d'image et une tâche de répétition de phrases, toutes deux extraites du protocole MonPaGe (Pernon et al., 2020). L'altération de ces paramètres est en effet caractéristique des troubles de la parole dans la DH (Teston & Viallet, 2005 ; Pinto & al., 2010). Néanmoins, il persiste un manque de consensus quant à leur évolution au cours de la maladie.

Les productions du patient avant et après PEC ont été comparées entre elles ainsi qu'aux productions de deux groupes contrôles : un groupe de 10 hommes sains (45 à 75 ans) et un groupe de 6 hommes MP (57 à 81 ans, stades 2 à 3 [3]), aucun n'ayant reçu d'entraînement par NMT.

L'outil de screening MonPaGe 2.0.s (Trouville et al., 2021) révèle une atténuation de l'altération de la parole passant de 2,5 (altération modérée à sévère ; cut-off >2) à 0 (absence d'altération). Les mesures sur la QoL et aspects thymiques n'indiquent pas d'évolution. Parmi les deux indices acoustiques étudiés, seule la variabilité de la F<sub>0</sub> semble impactée. En effet, dans les deux types de tâches, elle tend à décroître après entraînement et se rapproche des valeurs du groupe contrôle sain. Des analyses complémentaires sont en cours de réalisation.

### Références bibliographiques

- Goetz, C. G. (2010). Movement Disorder Society-Unified Parkinson's Disease Rating Scale (MDS-UPDRS) : une nouvelle échelle pour l'évaluation de la maladie de Parkinson. *Revue Neurologique*, 166(1), 1-4. <https://doi.org/10.1016/j.neurol.2009.09.001>
- Grand, M. (2019). *Manuel clinique de rééducation par la musique*. Louvain-la-Neuve : De Boeck Supérieur
- Koshimori, Y., & Thaut, M. H. (2018). Future perspectives on neural mechanisms underlying rhythm and music based neurorehabilitation in Parkinson's disease. *Ageing Research Reviews*, 47, 133-139. <https://doi.org/10.1016/j.arr.2018.07.001>
- Pernon, M., Levêque, N., Delvaux, V., Assal, F., Borel, S., Fougeron, C., Trouville, R., Laganaro, M. (2020). MonPaGe, un outil de screening francophone informatisé d'évaluation perceptive et acoustique des troubles moteurs de la parole (dysarthries, apraxie de la parole). *Rééducation Orthophonique*, 281, 171–197.
- Pinto, S., Ghio, A., Teston, B., Viallet, F. (2010). La dysarthrie au cours de la maladie de Parkinson. Histoire naturelle de ses composantes : dysphonie, dysprosodie et dysarthrie. *Revue neurologique*, 166, 800-810.
- Teston, B., Viallet, F. (2005). La dysprosodie parkinsonienne, in Ozsancak, C. *Les troubles de la parole et de la déglutition dans la maladie de Parkinson*. Solal, 161-193.
- Trouville, R., Delvaux, V., Fougeron, C., Laganaro, M. (2021). *Logiciel d'évaluation de la parole (version screening) MonPaGe-2.0.s* [Computer program].

# Fiabilité des valeurs de temps maximal de phonation, débit diadococinétique et débit de parole au travers de multiples essais par des participants neuro-typiques

Cécile FOUGERON<sup>1</sup>  
Nicolas AUDIBERT<sup>1</sup>

<sup>1</sup> Laboratoire de Phonétique et Phonologie,  
UMR 7018 CNRS,  
Université Sorbonne Nouvelle

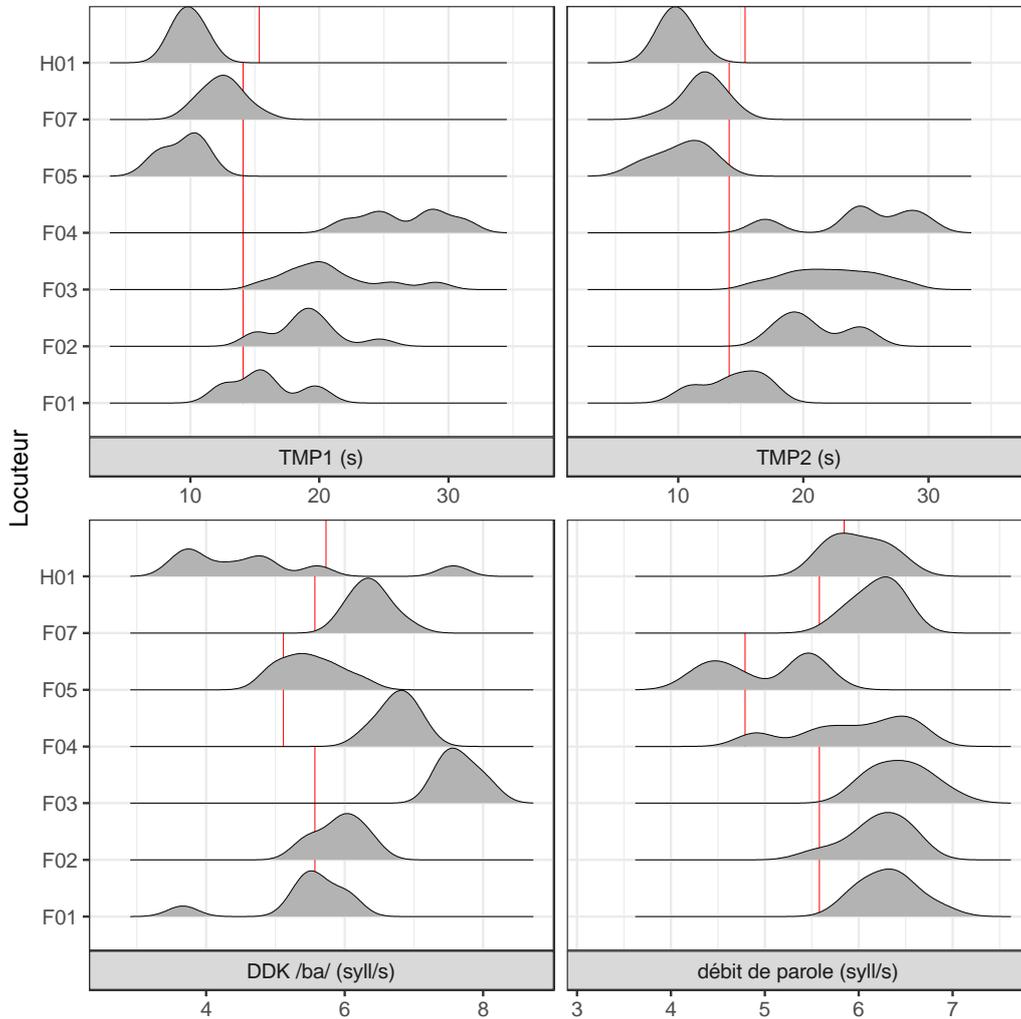
Les recherches ainsi que les pratiques d'évaluation et de remédiation en milieu clinique font appel à une grande variété de tâches de production, allant de la production de voyelles tenues au discours conversationnel. Une question fréquemment posée est de savoir si toutes ces tâches impliquent les mêmes mécanismes moteurs et cognitifs. Les tâches de diadococinésies (DDK), traditionnellement utilisées pour l'évaluation des troubles moteurs de la parole, sont particulièrement controversées (par ex. Maas, 2016 ; Ziegler, 2003a, 2003b). Si certains excluent ces tâches considérées comme « non speech », il s'avère difficile de les contourner car elles sont souvent sensibles et discriminantes, même avec des troubles peu sévères.

En effet, l'intérêt de ces tâches, qu'on considère comme des tâches de « performance maximale », est de pousser le système de production vers ses limites (Kent et al., 1987). Or, on peut se demander aussi si, à partir du moment où ces tâches impliquent des stratégies de 'performance', leur sensibilité à la variation (ex. fatigue, motivation du participant, changements biologiques, ...) est différente de celle d'une tâche de parole plus 'naturelle'.

L'objectif du travail présenté ici est d'examiner la fiabilité des valeurs obtenues sur plusieurs tâches de production en termes de reproductibilité au travers de multiples essais par les mêmes individus. Cette reproductibilité a été examinée pour les aspects liés à la qualité de voix ou la hauteur de  $f_0$ , mais sur peu d'autres dimensions (par ex. van Mersbergen et al., 1999 ; Vogel et al., 1991). Ici nous nous concentrons sur des aspects liés à la gestion temporelle : (a) le temps maximal de phonation sur un /a/ tenu (TMP) sur deux répétitions successives, (b) le débit diadococinétique sur la répétition de syllabes identiques (/bababa/, /gogogo/) et séquentielles (/badego/), et (c) le débit de parole obtenu sur la lecture d'une phrase.

Neuf locuteurs francophones ont été enregistrés pour l'ensemble de ces tâches (parmi d'autres) sur 8 à 10 sessions étalées sur deux mois. Pour le TMP, 10 autres sessions espacées d'un an, ont été enregistrés pour 8 des 9 locuteurs. La variabilité des mesures entre sessions chez un même individu est comparée à la variabilité inter-locuteur observée sur une cohorte de plus de 70 locuteurs en fonction des tranches d'âges.

Les résultats préliminaires illustrés sur la *Figure 1* pour quelques-unes de ces mesures chez 7 locuteurs montrent que la stabilité des mesures temporelles entre sessions dépend des locuteurs et des tâches.



*Figure 1.* Distribution des durées de TMP (1<sup>ère</sup> et 2<sup>ème</sup> répétition), du débit DDK pour la répétition de la syllabe /ba/, et du débit de parole sur la lecture de phrase, chez 7 locuteurs enregistrés sur 8 à 10 sessions. En arrière-plan est représentée la distribution observée pour plus de 70 locuteurs de même sexe et tranche d'âge : du 1<sup>er</sup> au 99<sup>e</sup> centile en rose, du 5<sup>e</sup> au 95<sup>e</sup> centile en violet et du 10<sup>e</sup> au 90<sup>e</sup> centile en bleu, la médiane étant représentée par la ligne verticale rouge.

### Références bibliographiques

- Maas, E. (2016) Speech and nonspeech: What are we talking about? *International Journal of Speech-Language Pathology*, 19, 345-359.
- Kent, R. D., Kent, J. F., & Rosenbek, J. C. (1987). Maximum Performance Tests of Speech Production. *Journal of Speech and Hearing Disorders*, 52(4), 367. doi:10.1044/jshd.5204.367

- Van Mersbergen, M. R., Verdolini, K., Titze, I. R. (1999). Time-of-day effects on voice range profile performance in young, vocally untrained adult females. *Journal of Voice*, *13*, 518–528.
- Vogel, A. P., Fletcher, J., Snyder, P. J., Fredrickson, A., & Maruff, P. (2011). Reliability, Stability, and Sensitivity to Change and Impairment in Acoustic Measures of Timing and Frequency. *Journal of Voice*, *25*(2), 137–149. doi:10.1016/j.jvoice.2009.09.003
- Ziegler, W. (2003a) Speech motor control is task-specific: evidence from dysarthria and apraxia of speech. *Aphasiology*, *17*, 3-36.
- Ziegler, W. (2003b) To speak or not to speak: Distinctions between speech and non-speech motor control. *Aphasiology*, *17*, 99-105.



# Relation entre productions verbales et gestuelles chez deux enfants porteurs d'une trisomie 21 : une étude préliminaire

Lou BRUN<sup>1</sup>  
Christelle DODANE<sup>2</sup>  
Fabrice HIRSCH<sup>1</sup>

<sup>1</sup> PRAXILING UMR 6267 CNRS,  
Université Paul Valéry, Montpellier 3

<sup>2</sup> CLESTIA EA 7345,  
Université Sorbonne Nouvelle

La trisomie 21 (T21\*) retarde l'acquisition du langage et la cognition non verbale (Chapman & Hesketh, 2000). Les capacités musculaires permettant l'articulation sont généralement touchées et on note des déficits de réception et production de langage, indépendamment d'un retard cognitif (Cleland et al., 2010). L'atteinte généralisée implique d'importantes difficultés de communication. Néanmoins, les enfants T21 ont des répertoires gestuels plus larges (Singer Harris et al., 1997), la fréquence de leurs gestes est similaire voire plus élevée que chez les enfants typiques (Iverson et al., 2003). Le lien entre l'abondance de gestes et le retard langagier n'a pas encore été établi, il l'a été chez les enfants dysphasiques (Stefanini et al., 2007).

Afin de vérifier si la gestualité pourrait favoriser l'intercompréhension familiale, nous élaborons un corpus de données longitudinales, et avons filmé deux enfants T21 (enfant 1 à 31, 37 mois et enfant 2 à 37, 44 mois). La transcription et l'annotation des vidéos (20 minutes chacune) via CLAN (MacWhinney, 2021) ont permis l'analyse des interactions parents-enfants. Nous avons systématiquement défini l'initiateur de l'interaction, les moyens de communications de l'enfant (verbaux et/ou gestuels), le résultat de la production ("efficaces", impliquant une réaction de l'adulte, sinon "infructueuses"), s'il y avait attention conjointe (AC) ou <pas (c'est-à-dire une attention portée sur un même objet par l'adulte et l'enfant, Tomasello, 1988) et les éventuelles compensations non verbales du niveau langagier. Ont été également annotés les gestes emblématiques, manuels co-verbaux, les mouvements corporels (buste, tête, corps entier) et la cible des regards. Le taux gestuel (nombre de gestes par rapport au nombre de mots) a été calculé pour chaque transcription (Tellier, 2014). Les interactions se font majoritairement à l'initiative des enfants et sont adressées à l'adulte. L'AC accompagne toujours les productions enfantines "efficaces". L'enfant 1 utilise moins de mots mais tous les types de gestes (3 types de mouvements, gestes manuels co-verbaux et emblématiques) alors que l'enfant 2 utilise plus de mots, mais seulement certains mouvements (buste, tête)

et les gestes manuels co-verbaux. Il semblerait donc que l'enfant plus jeune et moins verbal ait recours à davantage de types de gestes. Ces résultats préliminaires seront complétés par un plus grand nombre de données et le calcul du rapport geste/verbe dans les productions "efficaces". Des annotations du niveau prosodique permettront d'approfondir les liens entre les différentes modalités chez ces enfants et de déterminer si la gestualité permet effectivement de compenser le niveau langagier.

### Références bibliographiques

- Bernstein Ratner, N., Rooney, B., & MacWhinney, B. (1996). Analysis of stuttering using CHILDES and CLAN. *Clinical Linguistics and Phonetics*, 10(3), 169-18
- Chapman, R. S., & Hesketh, L. J. (s. d.). (2000). *Behavioral Phenotype of individuals with Down Syndrome*. 14.
- Cleland, J., Wood, S., Hardcastle, W., Wishart, J., & Timmins, C. (2010). Relationship between speech, oromotor, language and cognitive abilities in children with Down's syndrome. *International Journal of Language & Communication Disorders*, 45(1), 83– 95. <https://doi.org/10.3109/13682820902745453>
- Iverson, J. M., Longobardi, E., & Caselli, M. C. (2003). Relationship between gestures and words in children with Down's syndrome and typically developing children in the early stages of communicative development. *International Journal of Language & Communication Disorders*, 38(2), 179–197. <https://doi.org/10.1080/13682820310000628>
- MacWhinney, B. (2021). *The CHAT Transcription Format*. <https://doi.org/10.21415/3mhn-0z89>
- Singer Harris, N. G., Bellugi, U., Bates, E., Jones, W., & Rossen, M. (1997). Contrasting profiles of language development in children with williams and down syndromes. *Developmental Neuropsychology*, 13(3),345-370. <https://doi.org/10.1080/87565649709540683>
- Stefanini, S., Caselli, M. C., & Volterra, V. (2007). Spoken and gestural production in a naming task by young children with Down syndrome. *Brain and Language*, 101(3), 208– 221. <https://doi.org/10.1016/j.bandl.2007.01.005>
- Tellier, M. (2014). Quelques orientations méthodologiques pour étudier la gestuelle dans des corpus spontanés et semi-contrôlés. *Discours*,15. <http://journals.openedition.org/discours/8917>
- Tomasello, M. (1988). The role of joint attention in early language development. *Language Sciences*, 11, 69-88.

# Prosodie et gestualité chez les adolescents avec un trouble du spectre autistique

Cwiosna ROQUES<sup>1</sup>  
Christelle DODANE<sup>2</sup>  
Fabrice HIRSCH<sup>1</sup>

<sup>1</sup>Laboratoire PRAXILING,  
UMR 5267 CNRS,  
Université Paul Valéry Montpellier 3

<sup>2</sup>Laboratoire CLESTHIA,  
EA 7345,  
Université Sorbonne Nouvelle Paris 3

Le trouble du spectre autistique (TSA) est un trouble du neurodéveloppement caractérisé par un ensemble de déficiences dans la communication et l'interaction sociale (DSM-V, 2013). Notre travail se concentre sur deux dimensions de la possible utilisation atypique du langage dans l'autisme : la prosodie et le geste. Pris séparément, la prosodie et le geste ont reçu une certaine attention dans la littérature sur les TSA. En ce qui concerne la prosodie, les chercheurs ont noté des anomalies au niveau du rythme, de l'accentuation (absente, exagérée ou inappropriée), du ton (atone ou trop chantant), de l'intensité (cri ou chuchotement) et du débit de parole (trop rapide ou trop lent) (Baltaxe & Simmons, 1985 ; Globerson et al., 2015 ; Loveall et al., 2021). Pour les gestes, les TSA peuvent s'accompagner de déficits dans les comportements de communication non verbale, tels que des déficits dans le contact visuel et le langage corporel ou des déficits dans la compréhension et l'utilisation des gestes (Lelord, Sauvage, 1991 ; De Marchena et al., 2018).

Dans ce contexte, notre étude est bâtie sur l'analyse de la prosodie et des gestes afin d'identifier les spécificités potentielles liées à la communication des personnes autistes. Les individus avec TSA (Silverman et al., 2017) produisent les mêmes types de gestes coverbaux (iconique, déictique, battement, métaphorique, emblèmes) que leurs pairs neurotypiques, mais ces gestes sont moins nombreux et plus difficiles à comprendre, ce qui a des implications significatives sur leurs capacités de communication et d'intégration sociale (Silverman et al., 2017). Bien que la prosodie et les gestes coverbaux remplissent des fonctions pragmatiques liées à l'expressivité du discours (Bolinger, 1983 ; Level, 1989 ; Clark, 1994 ; Guellai et al., 2014), leur usage et leurs caractéristiques restent à ce jour méconnus. Comblé cette lacune est essentiel puisque ces aspects ont un impact significatif sur les plans linguistique et discursif, mais aussi sur la transmission des attitudes et des affects.

Pour ce faire, nous avons annoté la prosodie et les gestes de 5 enfants âgés de 10 à 15 ans filmés lors de deux tâches du programme d'observation diagnostique de l'autisme (ADOS-2, module 3, Lord et al., 2000) au CHU Lapeyronie de Montpellier (CHU Montpellier, cohorte ELENA ; Baghdadli et al., 2020). La première tâche consiste à raconter puis mimer une histoire à partir d'images, la seconde est un extrait de parole libre sur la thématique de l'amitié. La comparaison des deux tâches permet d'observer si les atypicités fréquemment rapportées dans les études sur les TSA se confirment dans nos données dans deux contextes différents :

celui d'une activité incitant à adapter la communication pour raconter une histoire et un contexte d'échange spontané. Pour la prosodie, ont été mesurés des paramètres d'intensité (intensité moyenne en décibels et écart type) et temporels (vitesse articulatoire). Les gestes des mains ont été classés comme représentatifs (déictiques, iconiques ou métaphoriques) ou non-représentatifs (illustrant la structuration du discours; McNeill, 1992; Kita & Özyürek, 2003), afin de vérifier si les gestes coverbaux diffèrent selon les activités. Nous présenterons nos principaux résultats et mettrons l'accent sur quelques exemples.

### Références bibliographiques

- American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders: DSM-V. *American Psychiatric Association*.
- Baghdadli A., Loubersac, J., Michelon, C., Rattaz, C., Pernon, E., Traver, S., Mortamais, M., Miot, S., Picot, M.C., & consortium Elena (2020). Cohorte Elena : étude transdisciplinaire des déterminants des troubles du spectre de l'autisme. BEH. [http://beh.santepubliquefrance.fr/beh/2020/6-7/2020\\_6-7\\_4](http://beh.santepubliquefrance.fr/beh/2020/6-7/2020_6-7_4).
- Baltaxe, C, & Simmons, J. Q. (1985). Prosodic development in normal and autistic children. In Schopler, E. & Mesibov, G.B. (eds.), *Communication problems in autism*. New-York: Plenum Press, 95-125.
- Bolinger, D. (1983). Intonation and gesture. *American Speech*, 58(2), 156-174.
- Clark, H.H. (1994). Managing problems in speaking. *Speech Communication*, 15(3-4), 243-250.
- Cruttenden, A. (1997). *Intonation* (2nd ed.) Cambridge: Cambridge University Press.
- Globerson, E., Amir, N., Kishon-Rabin, L., & Golan, O. (2015). Prosody recognition in adults with high-functioning autism spectrum disorders: From psychoacoustics to cognition: Prosody recognition and psychoacoustics in ASD. *Autism Research*, 8(2), 153–163.
- Guellai, B., Streri, A., & Yeung, H. H. (2014). The development of sensorimotor influences in the audiovisual speech domain: Some critical questions. *Frontiers in Psychology*, 5, 812.
- Kita, S., & Özyürek, A. (2003). What does cross-linguistic variation in semantic coordination of speech and gesture reveal?: Evidence for an interface representation of spatial thinking and speaking. *Journal Memory and Language*, 48(1), 16-32.
- Levelt, W.J. (1989). *Speaking: From intention to articulation*. Cambridge: MIT Press.
- Lord, C., Risi, S., Lambrecht, L., Cook, E. H., Leventhal, B. L., DiLavore, P. C., Pickles, A., & Rutter, M. (2000). The autism diagnostic observation schedule—Generic: A standard measure of social and communication deficits associated with the spectrum of autism. *Journal of Autism and Developmental Disorders*, 30(3), 205–223.
- Loveall, S. J., Hawthorne, K., & Gaines, M. (2021). A meta-analysis of prosody in autism, Williams syndrome, and Down syndrome. *Journal of Communication Disorders*, 89, 106055.
- MacWhinney, B. (2000). The CHILDES project: tools for analyzing talk. *Child language teaching and therapy*, 8(2).

- McNeill, D. (1992). *Hand and Mind: What gestures reveal about thought*. Chicago : The University of Chicago Press.
- McNeill, D. (2005). *Gesture & thought*. Chicago: The University of Chicago Press.
- Silverman, L.B., Eigsti, I.-M. & Bennetto, L. (2017). I tawt i taw a puddy tat: Gestures in canary row narrations by high-functioning youth with autism spectrum disorder. *Autism Research*, 10, 1353-1363.
- Simmons, J.Q., & Baltaxe, C. (1975). Language patterns of adolescent autistics. *Journal of Autism and Childhood Schizophrenia*, 5, 333-351.



# Modélisation individuelle de juges pour la prédiction automatique de l'intelligibilité de la parole

Sebastião QUINTAS<sup>1</sup>  
Julie MAUCLAIR<sup>1</sup>  
Julien PINQUIER<sup>1</sup>

<sup>1</sup> IRIT, Université de Toulouse, CNRS, Toulouse INP, UT3, Toulouse, France

Le développement d'outils automatiques pour la prédiction de l'intelligibilité de la parole peut être vu comme une alternative plus robuste, reproductible et objective face aux évaluations perceptives réalisées en clinique. En effet, ces évaluations peuvent être conditionnées par une variété d'aspects, tels que la connaissance préalable du patient ou la tâche d'évaluation de la parole elle-même (Balaguer et al., 2019). Les travaux de (De Bodt et al., 2002) ont montré que l'intelligibilité de la parole est liée à quatre paramètres perceptifs : Voix (V), Résonance (R), Prosodie (P) et Distorsions Phonémiques (PD), mais avec des niveaux de pertinence variables en fonction de chaque juge. Vu que l'intelligibilité est une mesure intrinsèquement subjective, la modélisation des juges individuels devient pertinente dans le contexte de l'intelligibilité de la parole (Woisard et al., 2021).

Dans cette étude, nous faisons la modélisation d'un ensemble de juges perceptifs, utilisés pour la prédiction de l'intelligibilité dans la cadre du corpus français du cancer des voies aérodigestives supérieures (Woisard et al., 2020). Nous nous basons sur le paradigme des *X-vector speaker embeddings* (Quintas et al., 2020) afin de proposer la modélisation illustrée sur la *Figure 1*. Chaque modèle individuel reçoit en entrée une représentation, qui est ensuite transmise à un réseau neuronal peu profond. Ce réseau prédit les quatre paramètres perceptifs : Voix (V), Résonance (R), Prosodie (P) et Distorsions Phonémiques (PD). Ensuite, l'intelligibilité est prédite en fonction de ces quatre paramètres, en utilisant une moyenne simple (poids uniformes) ou pondérée, qui utilise des poids différents pour chaque paramètre. Pour chercher les poids optimaux pour la moyenne pondérée, une méthode par grille de recherche (*grid search*) a été conçue.

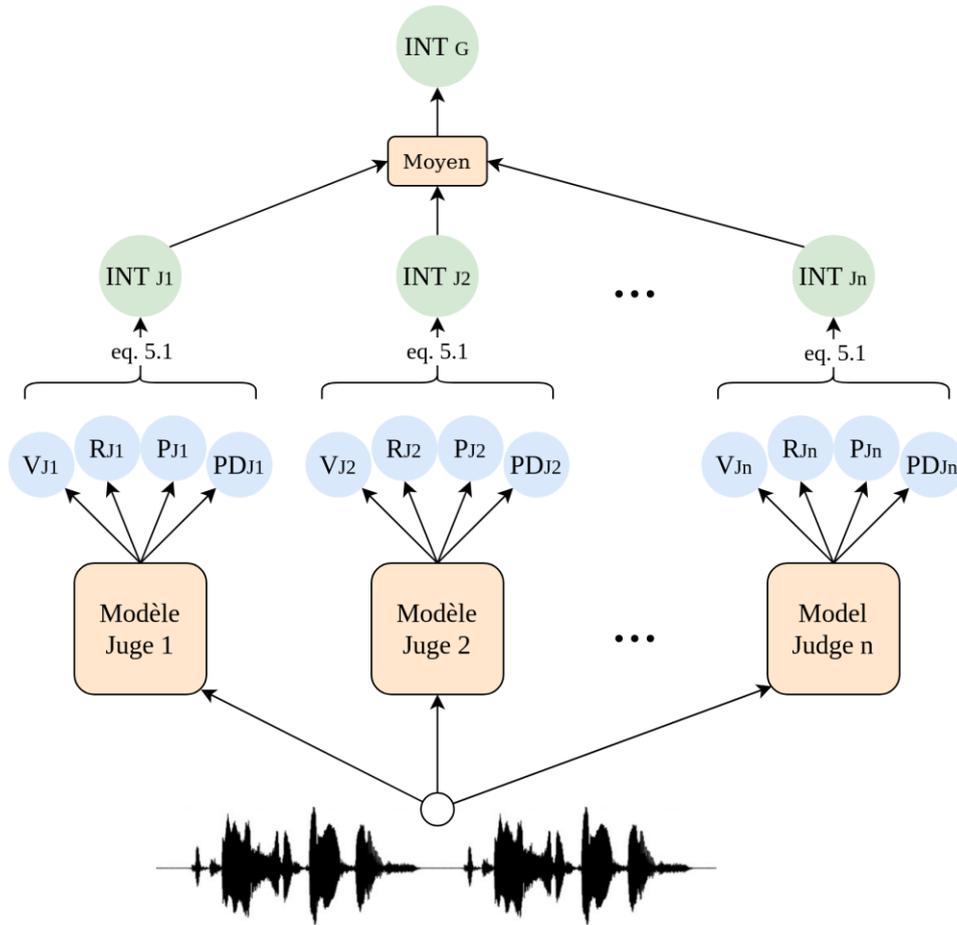


Figure 1. Schéma de la modélisation individuelle de chaque juge perceptif.

Dans le *Tableau 1*, nous pouvons observer deux analyses distinctes. La première montre la prédiction d'intelligibilité en utilisant les quatre paramètres perceptifs de chaque juge. La seconde montre exactement la même prédiction avec la modélisation automatique présentée précédemment. Malgré les corrélations inférieures dans la modélisation automatique de chaque juge, la différence entre les résultats obtenus avec les poids uniformes et les poids *grid search* est plus faible que dans le cas de l'analyse perceptive. Puisqu'il existe différents profils parmi les différents juges perceptifs (De Bodt et al., 2002), nous pouvons dire que la modélisation automatique favorise une prédiction, non seulement plus homogène entre les différents paramètres utilisés, mais également plus explicable. L'explicabilité des systèmes automatiques et des mesures produites devrait permettre une meilleure acceptabilité par les cliniciens.

Tableau 1. Corrélations obtenues en utilisant des moyennes des 4 paramètres perceptifs, pour l'analyse perceptive et la modélisation automatique des juges. Les modèles automatiques étaient entraînés avec les scores de référence de chaque juge en utilisant une validation croisée à 10 plis.

Juge	Analyse perceptive des juges		Modélisation automatique des juges	
	Poids Uniformes	Poids <i>Grid Search</i>	Poids Uniformes	Poids <i>Grid Search</i>
1	0,828	0,851	0,613	0,630
2	0,774	0,852	0,699	0,728
3	0,836	0,879	0,613	0,633
4	0,767	0,863	0,726	0,742
5	0,648	0,784	0,490	0,539
6	0,750	0,810	0,637	0,649

### Références bibliographiques

- Balaguer, M., Pommée, T., Farinas, J., Pinquier, J., Woisard, V., & Speyer, R. (2019). Effects of oral and oropharyngeal cancer on speech intelligibility using acoustic analysis : Systematic review. *Head & neck*, 42(1), 111-130. <https://doi.org/10.1002/hed.25949>
- De Bodt, M., Huici, M. E. H., & Van De Heyning, P. (2002). Intelligibility as a linear combination of dimensions in dysarthric speech. *Journal of Communication Disorders*, 35(3), 283-292. [https://doi.org/10.1016/s0021-9924\(02\)00065-5](https://doi.org/10.1016/s0021-9924(02)00065-5)
- Quintas, S., Mauclair, J., Woisard, V., & Pinquier, J. (2020). Automatic Prediction of Speech Intelligibility Based on X-Vectors in the Context of Head and Neck Cancer, *Proceedings of Interspeech*.
- Woisard, V., Balaguer, M., Fredouille, C., Farinas, J., Ghio, A., Lalain, M., Puech, M., Astésano, C., Pinquier, J., & Lepage, B. (2021). Construction of an automatic score for the evaluation of speech disorders among patients treated for a cancer of the oral cavity or the oropharynx: The Carcinologic Speech Severity Index. *Head & neck*, 44(1), 71-88. <https://doi.org/10.1002/hed.26903>
- Woisard, V., Astésano, C., Balaguer, M., Farinas, J., Fredouille, C., Gaillard, P., Ghio, A., Giusti, L., Laaridh, I., Lalain, M., Lepage, B., Mauclair, J., Nocaudie, O., Pinquier, J., Pouchoulin, G., Puech, M., Robert, D., & Roger, V. (2021c). C2SI corpus : a database of speech disorder productions to assess intelligibility and quality of life in head and neck cancers. *Language Resources and Evaluation*, 55(1), 173-190. <https://doi.org/10.1007/s10579-020-09496-3>



# La voix de la résilience : suivi acoustique des trajectoires de guérison des patients post-traumatiques au cours des séances de ré-exposition en imagination

Nadia GUEROUAOU<sup>1,2\*</sup>

Paul MAUBLANC<sup>1,4\*</sup>

Jean-Julien AUCOUTURIER<sup>1\*</sup>

Guillaume VAIVA<sup>2,3\*</sup>

<sup>1</sup>FEMTO-ST Institute UMR 6174,  
CNRS/UBFC, Besançon

<sup>2</sup> CHU Lille

<sup>3</sup> Centre National de Ressources et de Resilience (CN2R), Lille

<sup>4</sup>Laboratoire de Recherches Intégratives en Neurosciences et Psychologie  
Cognitive, Université de Franche-Comté, Besançon

*\* Les auteurs remercient S. Vanboove, S. Dubem, A. Damarey, F. Warembourg (CHU Lille) pour leur implication dans l'inclusion de patients*

Le trouble de stress post-traumatique (TSPT) est une maladie psychiatrique dans laquelle la parole est d'une importance considérable, car sa prise en charge dominante consiste à demander aux patients de raconter à voix haute leur événement traumatique au cours de plusieurs séances successives dites de « ré-exposition en imagination ».

Même si cette forme de thérapie est le *gold standard* du domaine (Brunet et al., 2018), les parcours individuels de guérison des patients varient beaucoup : certains voient d'abord leurs symptômes empirer, d'autres ne répondent pas au traitement. Comprendre et documenter ces différentes trajectoires individuelles de résilience en cours de traitement est important à la fois pour améliorer la prise en charge des patients (car le taux d'abandon est fort, entre 17 %-33 % ; Bradley et al., 2005), et pour une meilleure compréhension psychopathologique de la maladie.

Notre hypothèse est que l'évolution de séance en séance des caractéristiques acoustiques moyennes (ex. pitch moyen) de la voix des patients pourrait refléter l'évolution de leur pathologie, car ces caractéristiques sont connues par ailleurs pour varier avec les émotions chez le sujet sain (Owren & Bachorowski, 2007).

Dans cette étude, nous avons enregistré de façon longitudinale la voix de 30 patients TSPT suivis à la Consultation Régionale du Psychotrauma (CHRU Lille) d'octobre 2019 à juin 2021, pendant qu'ils racontaient leur épisode traumatique pendant la thérapie de ré-exposition. Les enregistrements (durée moyenne: 5.1 min., M=5.0 par patient) ont été analysés avec PRAAT pour extraire les caractéristiques acoustiques moyennes de chaque

séance (*pitch*, *jitter*, *shimmer*, *HNR* ; au delà des caractéristiques acoustiques moyennes, des futures analyses pourraient également s'intéresser à la variabilité de la *f0* et aux contours intonatifs). Ces caractéristiques ont été mises en relation avec une mesure clinique des symptômes effectuée à chaque séance, avec le questionnaire PCL5 (*post-traumatique checklist* ; Weathers et al., 1993), qui mesure notamment la force des symptômes d'intrusion de souvenirs (cluster d'items B), d'évitement (C), de pensées et cognitions négatives (D) et d'hyperactivité (E).

Ces mesures répétées montrent que, au fur et à mesure de la diminution générale de la sévérité des symptômes TSPT de séance en séance, le *pitch* moyen de la voix des patients diminue (-2.1Hz/session,  $t(70.1)=-3.7$ ,  $p<.01$ ), et leur *jitter* augmente, quoique non significativement. Dans le détail cependant, ces deux caractéristiques vocales répondent de façon contrastée à l'évolution de différents sous-groupes de symptômes testés par la PCL5. Pour le *pitch*, les patients voyant leurs symptômes d'intrusion de souvenirs (cluster B) diminuer voient leur *pitch* diminuer (-25Hz sur la totalité des symptômes,  $t(68.5)=-2.4$ ,  $p=.016$ ), alors que ceux voyant une amélioration des symptômes de pensées négatives (cluster D) voient leur *pitch* augmenter (+23Hz,  $t(68.8)=2.1$ ,  $p=.038$ ). Pour le *jitter*, les patients voyant une amélioration de l'intrusion (B) voient leur *jitter* diminuer (-1.0%,  $t(78.9)=-2.9$ ,  $p=.005$ ), alors que ceux voyant une amélioration des symptômes d'hyperactivité (cluster E) voient leur *jitter* augmenter (+1.1 %,  $t(76.3)=2.7$ ,  $p=.009$ ).

Ainsi, au-delà d'être seulement un biomarqueur de la guérison globale des patients, les mesures acoustiques de la voix ouvrent la possibilité d'une discrimination fine et personnalisée des trajectoires individuelles de résilience.

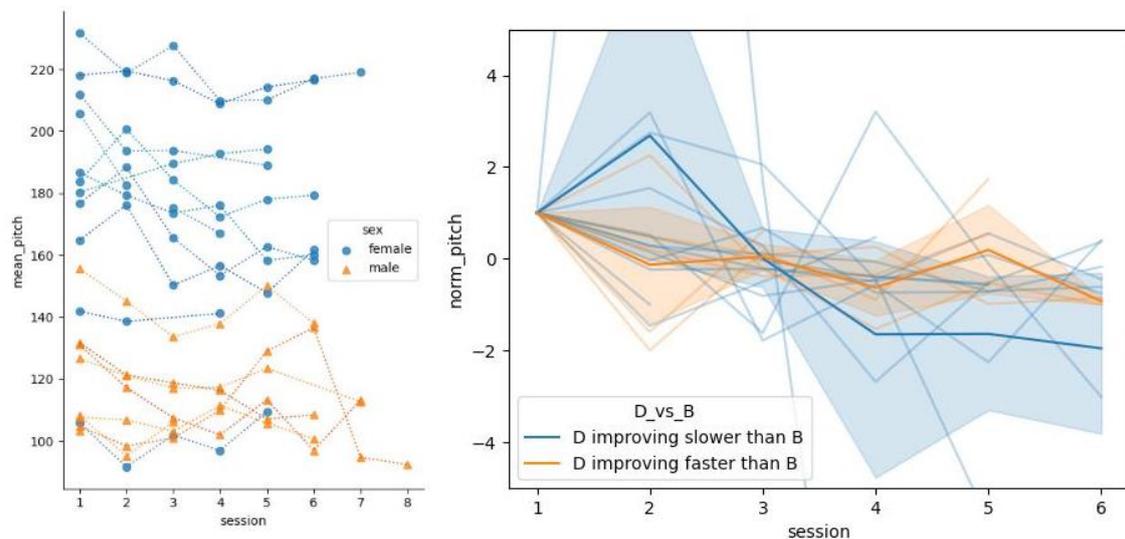


Figure 1. Évolution du pitch moyen des patients au cours de la thérapie. Même si le pitch des patients diminue au fur et à mesure de la diminution générale de la sévérité des symptômes TSPT de séance en séance (**gauche**), les patients dont les symptômes d'intrusion de souvenirs (« cluster B ») diminuent plus vite que les symptômes de pensées négatives (D) voient leur pitch diminuer plus fortement que ceux dont le parcours de guérison implique une diminution plus forte des symptômes D (**droite**).

## Références bibliographiques

- Bradley, R., Greene, J., Russ, E., Dutra, L., & Westen, D. (2005). A multidimensional meta-analysis of psychotherapy for PTSD. *American Journal of Psychiatry*, 162(2), 214–227.
- Brunet, A., Saumier, D., Liu, A., Streiner, D.L., Tremblay, J., & Pitman, R.K. (2018). Reduction of PTSD Symptoms with Pre-Reactivation Propranolol Therapy: A Randomized Controlled Trial. *The American journal of psychiatry*, 175(5), 427-433
- Owren, M., & Bachorowski, J. (2007). Measuring emotion-related vocal acoustics. *Handbook of emotion elicitation and assessment*, 239-266.
- Weathers, F., Litz, B., Herman, D., Huska, J., & Keane, T. (1993). *The PTSD Checklist (PCL): Reliability, Validity, and Diagnostic Utility*. Paper presented at the Annual Convention of the International Society for Traumatic Stress Studies, San Antonio, TX.



# Exploration de caractéristiques linguistiques et acoustiques pour la génération automatique de rapports de séances de remédiation cognitive avec un assistant virtuel

Yongxin ZHOU<sup>1</sup>  
Fabien RINGEVAL<sup>1</sup>  
François PORTET<sup>1</sup>

<sup>1</sup> Univ. Grenoble Alpes, CNRS, Inria, Grenoble INP, LIG,  
38000, Grenoble, France

Le projet THERADIA (Tarpin-Bernard et al., 2021) vise à créer un assistant virtuel accompagnant des personnes souffrant de troubles cognitifs lors de séances de remédiation réalisées à domicile. Le thérapeute délègue alors une partie de la gestion des séances à un assistant virtuel qui devient le pivot des interactions entre le patient et le thérapeute.

Une partie du projet s'intéresse à la mise en œuvre d'une génération automatique de rapports à destination des praticiens et résumant les séances de remédiation cognitive effectuées par les patients à domicile. Les informations collectées durant la séance sont traitées automatiquement pour en extraire des variables qui seront ensuite sélectionnées et structurées soit sous forme de texte, soit sous forme de tableau. Le traitement est basé sur un ensemble de règles expertes définies en collaboration avec des orthophonistes partenaires. À partir d'exemples de bilans orthophoniques, nous avons constitué le vocabulaire du domaine et identifié les éléments importants à intégrer dans un rapport. Le système de génération n'ayant pas vocation à faire du diagnostic, celui-ci ne rapporte que des observations descriptives, objectives et neutres, sans interprétation. Toutes les décisions sont prises par le thérapeute.

Cette analyse a permis de définir une taxonomie qui inclut des informations contextuelles, des éléments de communication langagière, ainsi que des aspects liés aux difficultés de compréhension.

Les éléments de communication langagière étudiés ont été définis à partir de travaux antérieurs portant sur l'évaluation automatique de troubles cognitifs à partir de caractéristiques linguistiques et acoustiques (Voletti et al., 2020). Nous avons exploité ces données, dont l'identification manuelle est à la fois difficile et fastidieuse, pour enrichir notre rapport.

Plus précisément, nous avons analysé des *caractéristiques lexicales*, qui incluent la taille du vocabulaire, ainsi que la diversité et la densité lexicale. Nous exploitons aussi des *caractéristiques syntaxiques*, estimées au niveau des énoncés (taille et durée moyenne). Enfin, une mesure du débit de parole vient compléter l'ensemble.

La *Figure 1* résume le processus du système de génération de résumé. Les informations multimodales sont capturées pendant la séance puis les variables d'intérêt sont extraites. Celles-ci sont ensuite structurées à l'aide de patrons linguistiques pour former des

paragraphes de documents préétablis. Les indicateurs linguistiques sont eux présentés sous forme de tableau.

La génération de rapports est une opportunité d'intégrer des mesures langagières dans la pratique de suivi de séances de remédiation cognitive, mesures qui demandent encore à être discutées avec la communauté des orthophonistes.

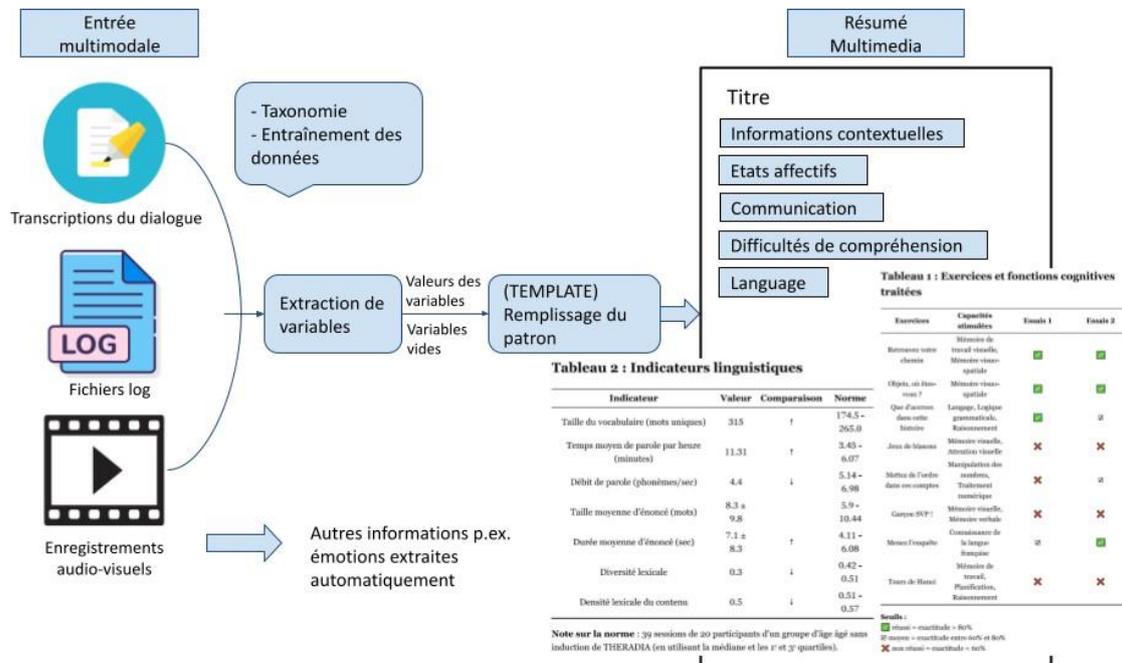


Figure 1. Illustration du système de génération à base de règles

## Remerciements

Cette recherche a été menée dans le cadre du projet THERADIA (financé par la BPI) et a été partiellement soutenue par MIAI@Grenoble-Alpes (ANR-19-P3IA-0003) et l'action COST Multi3Generation (CA18231).

## Références bibliographiques

- Tarpin-Bernard, F., Fruitet, J., Vigne, J., Constant, P., Chainay, H., Koenig, O., Ringeval, F., Bouchot, B., Bailly, G., Portet, F., Alisamir, S., Zhou, Y., Serre, J., Delerue, V., Fournier, H., Berenger, K. E., Zsoldos, I., Perrotin, O., Elisei, F., & Ghenassia, D. (2021). *THERADIA : Digital Therapies Augmented by Artificial Intelligence*. Springer eBooks, 478-485. [https://doi.org/10.1007/978-3-030-80285-1\\_55](https://doi.org/10.1007/978-3-030-80285-1_55)
- Voleti, R., Liss, J. M., & Berisha, V. (2020). A Review of Automated Speech and Language Features for Assessment of Cognitive and Thought Disorders. *IEEE Journal of Selected Topics in Signal Processing*, 14(2), 282-298. <https://doi.org/10.1109/jstsp.2019.2952087>

# Reconnaissance automatique des expressions émotionnelles dans la parole naturelle : le cas des patients dans des appels d'urgence à Besançon et à Lausanne

Dongjun WEI<sup>1</sup>  
Mohamed EMBARKI<sup>1</sup>  
Oussama BARAKAT<sup>2</sup>  
Coralie VAUCHERÉY<sup>1,2</sup>  
Thibaut DESMETTRES<sup>2,3</sup>  
Tania MARX<sup>2,4</sup>  
Stephan ROBERT<sup>5</sup>

<sup>1</sup>ELLIADD EA 4661,  
Université de Franche-Comté, Besançon

<sup>2</sup>LNIT EA 4662,  
Université de Franche-Comté, Besançon

<sup>3</sup>Hopitaux universitaires Genève

<sup>4</sup>Centre Hospitalier Régional Universitaire de Besançon

<sup>5</sup>Haute Ecole d'Ingénierie et de Gestion du Canton de Vaud

Tout changement d'état émotionnel s'accompagne d'une modulation des caractéristiques acoustiques (Scherer, 2003). Dans de nombreuses mesures acoustiques de la parole émotionnelle (Bänziger et al., 2001 ; Cowie et al., 2001), la capacité des caractéristiques des F0<sup>2</sup> à discriminer la parole émotionnelle a été généralement reconnue (Liebermann & al., 1962 ; Williams & Stevens, 1972 ; Mozziconacci & Hermes, 2000 ; Lee & al., 2014 ; Ali & al., 2015).

Cette étude se concentre sur les caractéristiques acoustiques des expressions émotionnelles dans la parole naturelle<sup>3</sup>, dans le but de construire des systèmes de reconnaissance automatique des émotions adaptés au domaine de la régulation médicale d'urgence.

L'échantillon de parole naturelle comprenait 32 appels téléphoniques de patients aux services des urgences du CHRU de Besançon et du CHUV de Lausanne. Les appels

---

<sup>2</sup> F0 moyenne, étendue de F0, formants sont souvent utilisés dans la reconnaissance des émotions. Dans notre recherche, les mesures de F0 ont ciblé les quatre points d'inflexion de chaque énoncé (début, fin, maximum et minimum) ainsi que la F0 moyenne.

<sup>3</sup> Beaucoup de ces études sur la parole émotionnelle sont basées sur la parole non naturelle (induite, simulée ou synthétisée)

téléphoniques exploités sont équilibrés en nombre<sup>4</sup> et en genre<sup>5</sup>. Seuls les extraits correspondant aux 4 émotions<sup>6</sup> à polarité négative ont été retenus. L'étiquetage des émotions prosodiques est largement basé sur les sentiments subjectifs des étiqueteurs, mais pour mieux distinguer les émotions des patients, une explication précise leur a été fournie préalablement.

Les extraits ont été d'abord segmentés et étiquetés ; ensuite leurs propriétés acoustiques sont mesurées manuellement<sup>7</sup> sous Praat (Boersma, 2017) ; finalement, 406 énoncés retenues correspondent aux émotions vocales non linguistiques. Les mesures de F0 ont ciblé les quatre points d'inflexion de chaque énoncé<sup>1</sup> et la F0 moyenne, il en est de même des mesures d'intensité. L'écart de F0 est converti en écart tonal. Les mesures de durée ont ciblé la vitesse d'articulation<sup>8</sup> et la vitesse de parole<sup>9</sup>.

Une comparaison globale des données montre que les indices acoustiques de l'émotion verbale varient très peu en intensité et en durée, mais que les indices F0 de l'émotion verbale pour les patients de genre différent dans les différents pays ne peuvent pas être résumés dans un modèle simple et fixe, e.g. les femmes des différents pays ont des F0 moyennes plus élevées pour l'angoisse et la tristesse et plus faibles pour l'embarras et la colère, tandis que les hommes ont des F0 plus élevées pour l'angoisse et la colère et plus faibles pour l'embarras et la tristesse...

La reconnaissance des émotions verbales naturelles est plus complexe que la reconnaissance des émotions verbales non naturelles. Les résultats de cette étude sont pour l'instant indicatifs, car la taille de l'échantillon est limitée, comparativement à la base de données qui contient plusieurs millions d'appels téléphoniques qui sont en cours de traitement.

---

<sup>4</sup> 16 appels de chaque hôpital

<sup>5</sup> 16 femmes et 16 hommes

<sup>6</sup> Angoisse, colère, embarras, tristesse

<sup>7</sup> En raison de la complexité de l'environnement d'enregistrement.

<sup>8</sup> Syllabes par seconde sans compter les pauses

<sup>9</sup> Syllabes par seconde, y compris les pauses

## Références bibliographiques

- Ali, S.A., Khan, A., & Bashir, N. (2015). Analyzing the Impact of Prosodic Feature (Pitch) on Learning Classifiers for Speech Emotion Corpus. *International Journal of Information Technology and Computer Science(IJTCS)*, 7(2), 54-59. <https://doi.org/10.5815/ijitcs.2015.02.07>
- Boersma, P. (2017). Praat, a system for doing phonetics by computer. *Glott International*, 5(9-10), 341-345.
- Bänziger, T., Grandjean, D., Bernard, P. J., Klasmeyer, G., & Scherer, K. R. (2001). Prosodie de l'émotion : étude de l'encodage et du décodage. *Cahiers de linguistique française*, 23, 11-37.
- Cowie, R., Douglas-Cowie, E., Tsapatsoulis, N., Votsis, G., Kollias, S., Fellenz, W., & Taylor, J. G. (2001). Emotion recognition in human-computer interaction. *IEEE Signal Processing Magazine*, 18(1), 32-80. <https://doi.org/10.1109/79.911197>.
- Lee, C., & Lui, S. (2014). Visualization of time-varying joint development of pitch and dynamics for speech emotion recognition. *The Journal of the Acoustical Society of America*, 135, 2422. <https://doi.org/10.1121/1.4878044>
- Lieberman, P., & Michaels, S. B. (1962). Some aspects of fundamental frequency and envelope amplitudes as related to the emotional content of speech. *Journal of the Acoustical Society of America*, 34(7), 922–927. <https://doi.org/10.1121/1.1918222>
- Mozziconacci, S. J. L., & Hermes, D. J. (2000). *Variations temporelles communiquant l'émotion dans la parole*. In Proceedings of the 23rd Journées d'étude sur la parole, JEP-2000, Aussois, France. 145-148.
- Scherer, K. R. (2003). Vocal communication of emotion: A review of research paradigms. *Speech Communication*, 40, 227–256.
- Williams, C. E., & Stevens, K. N. (1972). Emotions and speech: some acoustical correlates. *The Journal of the Acoustical Society of America*, 52(4), 1238–1250. <https://doi.org/10.1121/1.1913238>



# La polypose naso-sinusienne : impacts spectraux. Étude préliminaire.

Amélie ELMERICH<sup>1</sup>  
Angélique AMELOT<sup>1</sup>  
Lise CREVIER-BUCHMAN<sup>1,2</sup>  
Shinji MAEDA<sup>1</sup>  
Jean-François PAPON<sup>3</sup>

<sup>1</sup>Laboratoire de Phonétique et Phonologie,  
UMR7018 CNRS,  
Université Sorbonne-Nouvelle, Paris, France

<sup>2</sup>Hôpital Foch, Service de Laryngologie Phoniatrie  
Suresnes, France

<sup>3</sup>Hôpital Bicêtre, Service d'Oto-Rhino-Laryngologie,  
Le Kremlin Bicêtre, France

Certaines pathologies, comme la polypose naso-sinusienne primitive (PNS)<sup>10</sup>, perturbent la résonance nasale et la qualité des sons en obstruant les voies nasales et les sinus (Hongetal.,1997).L'énergie spectrale peut être altérée dans les basses fréquences. En effet, l'amplitude des premiers harmoniques est sensible à la nasalité (Dickson, 1962 ; Hattori et al., 1958 ; Simpson, 2012).

L'objectif de cette étude est d'évaluer l'impact de la chirurgie via une analyse spectrale des voyelles sur une cohorte de 4 patients atteints de PNS.

Des enregistrements acoustiques sont recueillis grâce à un masque en fibre de papier (Honda & Maeda, 2008) permettant d'enregistrer les débits oral et nasal tout en préservant l'acoustique (Elmerich et al., 2020). Des mesures spectrales sont effectuées au milieu des voyelles [a] dans un contexte VCV où V = [a] et C =[b, p, m, n, s, z, d, t, v] sur 4 hommes atteints d'une PNS avant et après chirurgie. Les coupes spectrales sont réalisées à la fois sur le signal acoustique du microphone, placé devant le masque à 5 cm de la bouche appelé signal acoustique oral (AKG C520L) et sur le signal dérivé du capteur de pression nasale, placé dans le compartiment nasal, appelé signal acoustique nasal. Les mesures sont obtenues sur 3 cycles en tenant compte de la fréquence du fondamental du locuteur. Nous avons mesuré le spectre moyen entre 0 et 2kHz ainsi que la différence entre les 4 premiers harmoniques exprimée en logarithme.

Nous nous focalisons ici sur le logatome [ada] pour le patient PM02 (*Figure 1*) en comparaison avec un témoin (MT03). Les résultats montrent une baisse du spectre moyen nasal (en bleu) et une augmentation du spectre moyen oral (en rouge) en postopératoire sur les voyelles du logatome. Les valeurs postopératoires se rapprochent des valeurs du locuteur témoin. La différence entre les 4 premiers harmoniques (numérotés sur les coupes spectrales de la *Figure*

---

<sup>10</sup> Inflammation chronique bilatérale des cavités naso-sinusiennes caractérisée par le développement de polypes. On parle de PNS primitive lorsqu'elle apparaît de manière isolée, c'est la forme la plus fréquente (2 à 4% de la population) (Coste, 2011).

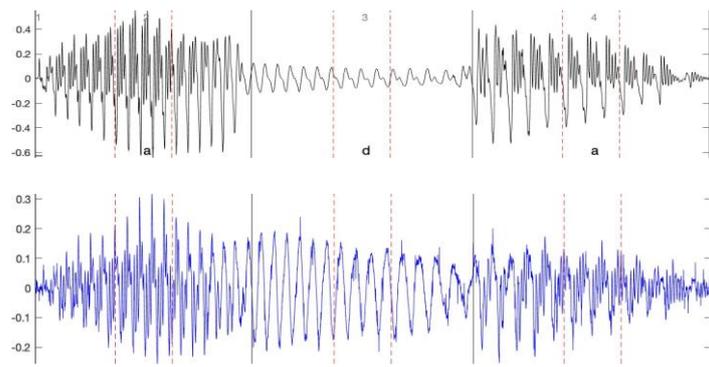
1) du spectre oral et nasal (H1-H1N, H2-H2N, *etc.*) augmente en postopératoire. On observe une plus grande dissociation entre les spectres oral et nasal. Les trois autres patients, quel que soit le contexte consonantique, suivent la même tendance. De même que pour les témoins qui ont des résultats similaires à MT03.

Tableau 1. Valeurs des spectres moyens oraux et nasals (dB) sur le logatome [ada]

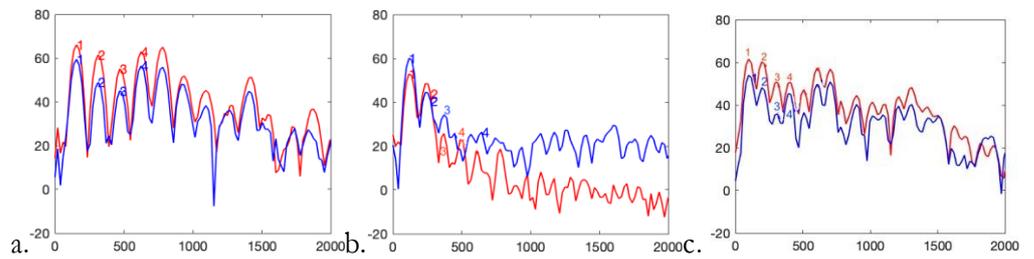
		Spectre moyen issu du signal acoustique nasal (dB)	Spectre moyen issu du signal acoustique oral (dB)
[a] (V1)	préopératoire	31,49	37,17
	postopératoire	25,92	41,50
	témoin	27,33	40,84
[d]	préopératoire	23,54	7,64
	postopératoire	20,75	19,96
	témoin	21,96	16,85
[a] (V2)	préopératoire	29,83	36,25
	postopératoire	25,94	39,08
	témoin	26,89	43,63

L'amplitude acoustique de la cavité orale apparaît plus importante en postopératoire en raison du désencombrement naso- sinusien. Après chirurgie, le ratio acoustique entre les cavités orale et nasale paraît meilleure. La chirurgie semble améliorer la résonance vocale.

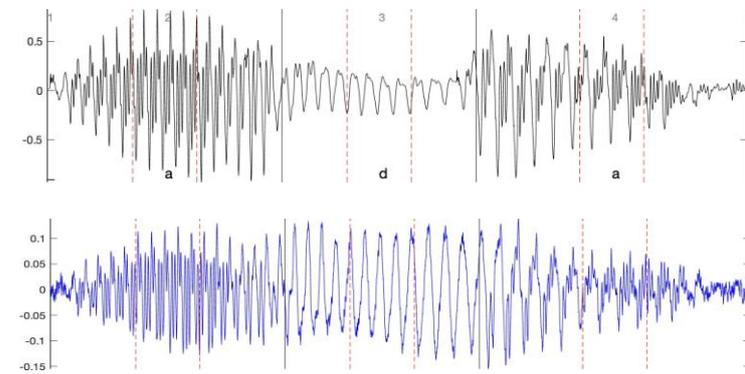
1. Avant chirurgie



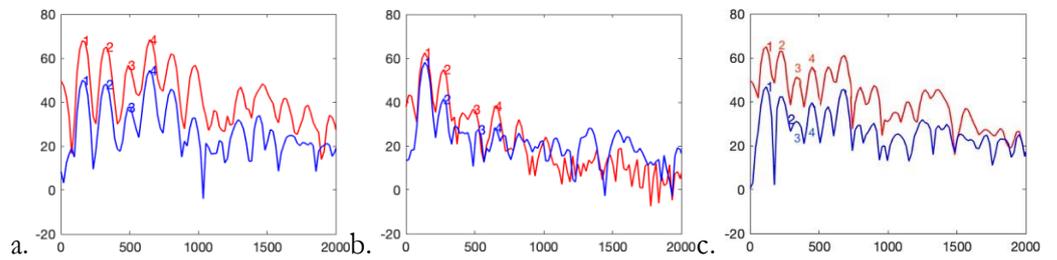
		V1	V2
H1	Oral	66	61
	Nasal	59	53
H2	Oral	61	60
	Nasal	49	48
H3	Oral	55	51
	Nasal	45	36
H4	Oral	63	50
	Nasal	56	39



2. Après chirurgie



		V1	V2
H1	Oral	68	65
	Nasal	50	47
H2	Oral	65	62
	Nasal	48	42
H3	Oral	57	51
	Nasal	38	29
H4	Oral	68	56
	Nasal	54	40



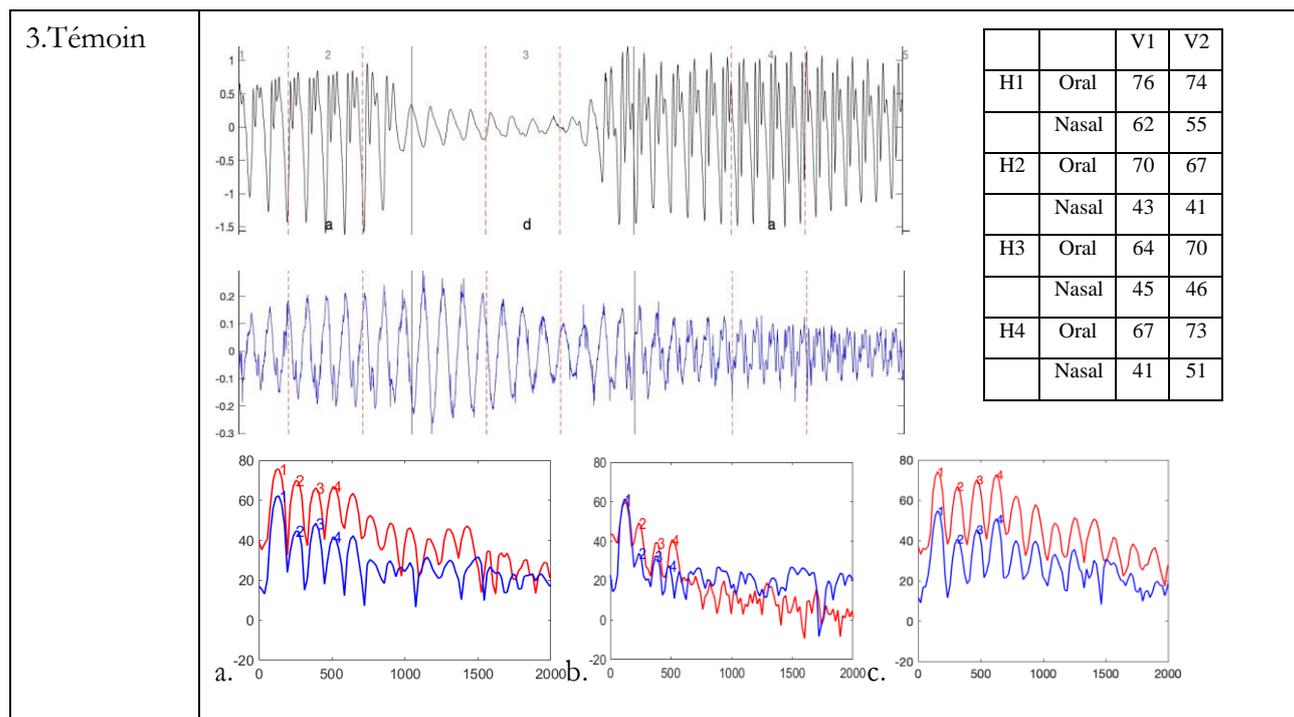


Figure 1. Du haut vers le bas : signal acoustique oral, nasal du logatome [ada] ainsi que les coupes spectrales oral (rouge) et nasal (bleu) sur [a] (V1) en a., [d] en b. et [a] (V2) en c., ainsi que sur la droite, un tableau reprenant les valeurs des harmoniques en dB sur le signal oral et nasal, avant chirurgie (1), après chirurgie (2) et pour un témoin (3).

### Références bibliographiques

- Coste, A. (2011). Polypose naso-sinusienne et médecine interne, *La Revue de médecine interne*, 32.
- Dickson, D. R. (1962). An Acoustic Study of Nasality. *Journal of Speech and Hearing Research*, 5(2), 103-111.
- Elmerich, A., Amelot, A., Maeda, S., Laprie, Y., Papon, J. F., & Crevier-Buchman, L. (2020). F1 and F2 measurements for French oral vowel with a new pneumotachograph mask. *ISSP 2020-12th International Seminar on Speech Production*.
- Hattori, S., Yamamoto, K., & Fujimura, O. (1958). Nasalization of Vowels in Relation to Nasals. *The Journal of the Acoustical Society of America*, 30(4), 267-274.
- Honda, K., & Maeda, S. (2008). Glottal-opening and airflow pattern during production of voiceless fricatives: a new non-invasive instrumentation. *The Journal of the Acoustical Society of America*, 123(5), 3738-3738.
- Hong, K. H., Kwon, S. H., & Jung, S. S. (1997). The assessment of nasality with nasometer and sound spectrography in patients with nasal polyposis. *Otolaryngology--Head and Neck Surgery*, 117(4), 343-348.
- Simpson, A. P. (2012). The first and second harmonics should not be used to measure breathiness in male and female voices. *Journal of Phonetics*, 40(3), 477-490.

# Différents stades d'Alzheimer : Rôle des méthodes automatisées de traitement de la prosodie et de la qualité de la voix

Jalal AL-TAMIMI<sup>1</sup>

<sup>1</sup> Laboratoire de Linguistique Formelle,  
UMR 7110 CNRS,  
Université Paris Cité

La détection précoce des signes du déclin cognitif (DC) associé à la maladie d'Alzheimer (Alz) permet potentiellement de ralentir la progression de la maladie (van Dyck et al., 2022). Généralement, des variations prosodiques (augmentation de pause, réduction du débit de parole) et de qualité de la voix (QV ; augmentation du souffle, réduction d'énergie spectrale) sont bien identifiées dans les stades avancés d'Alz (Ivanova et al., 2022; Shah et al., 2021). Cependant, la prévalence de ces caractéristiques n'est pas bien établie dans les stades légers, précoces (Déclin Cognitif Subjectif, DCS) ou préclinique (Déclin Cognitif Objectif, DCO).

Cette étude tente d'établir un profil prosodique et de QV des différents stades de la maladie. 119 participants espagnols/catalans (fondation « Alzheimer Centre Educativo (ACE) » à Barcelone) ont participé à cette étude (71 hommes ; 48 femmes ; cinq groupes : 19 contrôles neurotypiques, CN ; 31 DCS ; 16 naDCO non-amnésiques, nDCO et 23 amnésiques, aDCO ; 31 Alz légers ; tâches de production par participant : réponses aux questions dans le passé et le futur, description d'image, construction d'une scène).

Des analyses automatisées avec Prosogram (Mertens, 2004) ont été effectuées avec : 1) un profil prosodique avec des mesures temporelles et fréquentielles ( $f_0$ ) et 2) un profil de QV (version modifiée de Prosogram ; Shimmer, Jitter, énergie) suivant l'« Acoustic Breathiness Index » (Bartsties v. Latoszek et al., 2017). Des Régressions Multinomiales prédisant les risques associés entre stades et mesures acoustiques, montrent qu'en comparaison avec CN, la parole en Alz est associée à : réduction de quantité des productions, augmentation de pause et de  $f_0$ , augmentation du souffle et réduction d'énergie. L'aDCO se rapproche davantage d'Alz ; naDCO et DCS se rapproche de CN avec des différences marquées sur la QV ; profils variables permettant une discrimination entre les groupes (Voir *Tableau 1*).

En classification, seules les discriminations binaires permettent de valider les profils : les mesures neuropsychologiques montrent une discrimination parfaite (à 100%) entre CN ou DCS vs Alz ou naDCO ; discrimination proche des 80% avec la prosodie et/ou la QV. La QV a permis une meilleure discrimination entre CN vs DCS (82%), contre 50% avec les mesures neuropsychologiques ; l'ajout de la prosodie et/ou QV aux mesures neuropsychologiques a amélioré les taux (90% vs 80%).

Les mesures automatisées de la prosodie et de la QV permettent de profiler les stades de DC à risque du développement d'Alz et de les discriminer avec des taux proches de ceux des mesures neuropsychologiques.

*Tableau 1.* Résumé des patterns observés dans les domaines prosodiques et de qualité de la voix (QV). ↑ = augmentation ; ↓ = baisse ; → = différence minimale. Énergie = Énergie dans les hautes fréquences.

Domaine	Type	CN	DCS	nDCO	naDCO	Alz
Prosodie	Quantité des productions	→	↑	↓→	↓	↓
	Pause	→	↓	↑→	↑	↑
	F0	→	↓	↓→	↑	↑
QV	Shimmer	→	↓	↓→	↑	↑
	Jitter	→	↑	↓→	↑	↑
	Énergie	→	↑	↑→	↓	↓

### Références bibliographiques

- Barsties v. Latoszek, B., Maryn, Y., Gerrits, E., & De Bodt, M. (2017). The Acoustic Breathiness Index (ABI): A Multivariate Acoustic Model for Breathiness. *Journal of Voice*, 31(4), 511.e11-511.e27. <https://doi.org/10.1016/j.jvoice.2016.11.017>
- Ivanova, O., Meilán, J. J. G., Martínez-Sánchez, F., Martínez-Nicolás, I., Llorente, T. E., & González, N. C. (2022). Discriminating speech traits of Alzheimer's disease assessed through a corpus of reading task for Spanish language. *Computer Speech & Language*, 73, 101341. <https://doi.org/10.1016/j.csl.2021.101341>
- Mertens, P. (2004). The prosogram: Semi-automatic transcription of prosody based on a tonal perception model. *Paper Presented at the Speech Prosody*
- Shah, Z., Sawalha, J., Tasnim, M., Qi, S., Stroulia, E., & Greiner, R. (2021). Learning Language and Acoustic Models for Identifying Alzheimer's Dementia From Speech. *Frontiers in Computer Science*, 3. <https://www.frontiersin.org/article/10.3389/fcomp.2021.624659>
- van Dyck, C. H., Swanson, C. J., Aisen, P., Bateman, R. J., Chen, C., Gee, M., Kanekiyo, M., Li, D., Reyderman, L., Cohen, S., Froelich, L., Katayama, S., Sabbagh, M., Vellas, B., Watson, D., Dhadda, S., Irizarry, M., Kramer, L. D., & Iwatsubo, T. (2022). Lecanemab in Early Alzheimer's Disease. *New England Journal of Medicine*, NEJMoa2212948. <https://doi.org/10.1056/NEJMoa2212948>

# Etude des signaux vibroacoustiques de déglutition chez les sujets sains

Lila GRAVELLIER<sup>1,2</sup>  
Maxime LE COZ<sup>2</sup>  
Jérôme FARINAS<sup>1</sup>  
Fabrice NEVEU<sup>2</sup>  
Julien PINQUIER<sup>1</sup>

<sup>1</sup>IRIT, Université de Toulouse, CNRS, Toulouse INP, UT3,  
Toulouse, France

<sup>2</sup> Swallis Medical,  
Toulouse, France

La déglutition est un mécanisme physiologique essentiel permettant le passage des aliments de la bouche à l'estomac, tout en protégeant les voies aériennes. Plusieurs dysfonctionnements peuvent apparaître tout au long de cette chaîne menant à des troubles de la déglutition, aussi appelés *dysphagie*.

La dysphagie affecte 8% de la population mondiale (Cichero et al., 2013). Ces troubles peuvent altérer la qualité de vie du patient voire aggraver son pronostic. Les principales conséquences sont la dénutrition et la pneumonie par aspiration qui peut entraîner des troubles respiratoires graves, allant parfois jusqu'au décès. C'est pourquoi il est important de diagnostiquer les patients affectés le plus tôt possible. Cependant, les explorations cliniques actuelles restent subjectives, et les examens de référence invasifs, coûteux et peu accessibles.

L'auscultation cervicale à haute résolution est une alternative prometteuse pour l'évaluation et le suivi de la dysphagie. Elle consiste à placer des capteurs sur le cou du patient et à analyser les signaux de déglutitions enregistrés. Plusieurs travaux académiques ont montré l'intérêt de certains capteurs pour évaluer la déglutition, tels que le microphone et l'accéléromètre (Donohue et al., 2020; Dudik et al., 2015, 2018). Mais à ce jour, aucun n'a abouti à la création d'un dispositif utilisable en clinique présentant ces deux capteurs. Par ailleurs, les études s'intéressant aux signaux vibroacoustiques de déglutition de la population saine sont encore trop peu nombreuses et ne donnent pas forcément toutes les informations nécessaires à la comparaison ou la reproductibilité des résultats obtenus (So et al., 2022).

Ainsi, un protocole d'enregistrement détaillé a été établi pour étudier les signaux vibroacoustiques de déglutition chez la population saine en utilisant un nouveau dispositif d'auscultation cervicale à haute résolution : le Swallis DSA<sup>TM</sup>, présenté dans la *Figure 1*. Les nombreuses tâches qui le composent permettent d'obtenir des déglutitions contrôlées et spontanées, sur plusieurs textures et volumes, mais aussi d'autres activités pharyngolaryngées telles que la toux ou la parole (voir *Figure 2*). 97 sujets sains, âgés de 21 à 89 ans ont été enregistrés en suivant ce protocole.

Les premières analyses des données ont permis de valider l'intérêt de ces capteurs et le bon fonctionnement du dispositif. Les durées moyennes des phases détectées sur 884 déglutitions contrôlées correspondent bien aux références de la littérature (Morinière et al., 2008). Dans de futurs travaux, ce corpus permettra de mieux décrire et identifier les déglutitions en situation écologique à travers l'accéléromètre et le microphone.

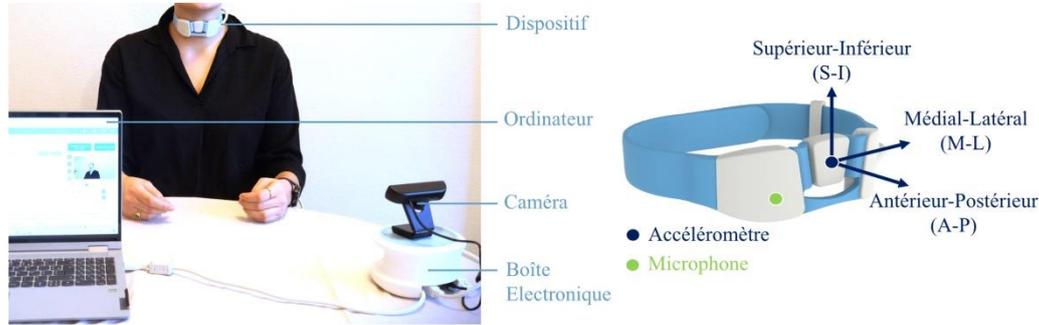


Figure 1: Mise en place du dispositif Swallis DSA™ pour l'enregistrement du protocole et positions des capteurs sur le collier. Sur le cou du patient, l'accéléromètre trois axes se place sur la partie inférieure de la proéminence laryngée et le microphone au niveau du cartilage cricoïde.

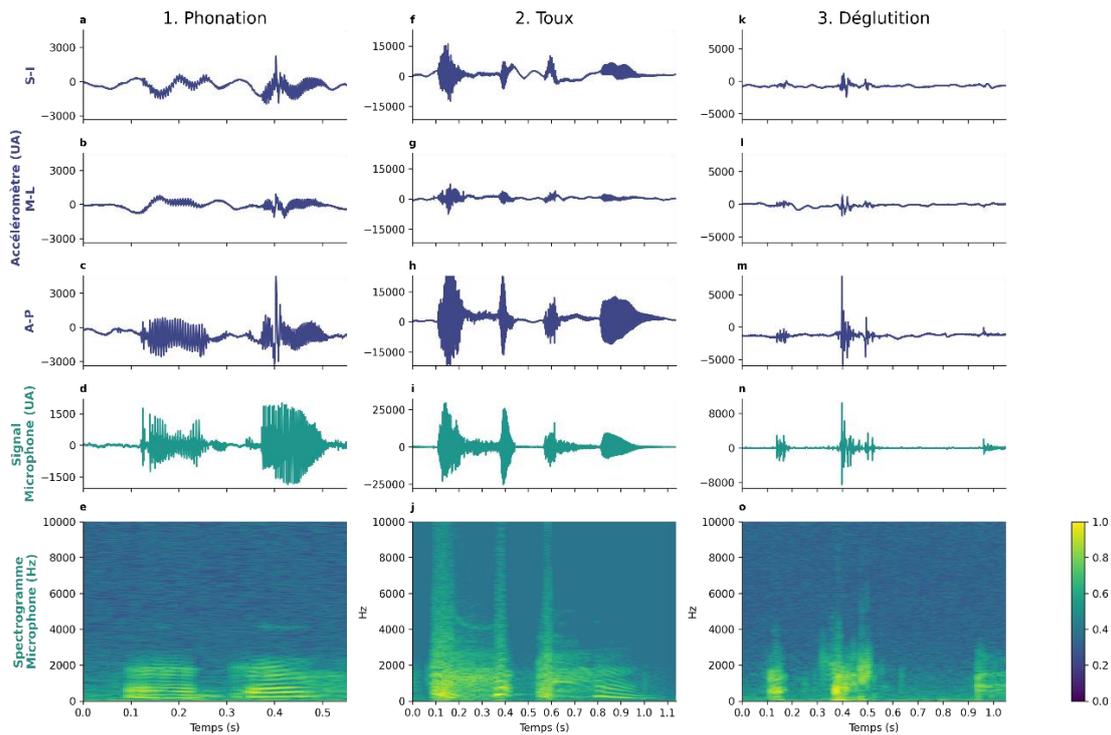


Figure 2 : Signaux vibroacoustiques de trois activités pharyngolaryngées : une phonation, une toux et une déglutition. De haut en bas, les trois axes de l'accéléromètre (Supérieur-Inférieur S-I, Médial-Latéral M-L et Antérieur-Postérieur A-P), le signal du microphone (représentés en unité arbitraire UA) et le spectrogramme du microphone. Ces activités génèrent principalement des vibrations perpendiculaires à la surface du cou (direction A-P).

Dans le cas de la déglutition, les mouvements d'antériorisation et d'élévation du larynx ainsi que les vibrations liées au passage du bolus vont produire du son et des vibrations détectés par les capteurs.

### Références bibliographiques

- Cichero, J. A. Y., Steele, C., Duivesteyn, J., Clavé, P., Chen, J., Kayashita, J., Dantas, R., Lecko, C., Speyer, R., Lam, P., & Murray, J. (2013). The Need for International Terminology and Definitions for Texture-Modified Foods and Thickened Liquids Used in Dysphagia Management : Foundations of a Global Initiative. *Current Physical Medicine and Rehabilitation Reports*, 1(4), 280-291. <https://doi.org/10.1007/s40141-013-0024-z>
- Donohue, C., Khalifa, Y., Perera, S., Sejdić, E., & Coyle, J. L. (2020). A Preliminary Investigation of Whether HRCA Signals Can Differentiate Between Swallows from Healthy People and Swallows from People with Neurodegenerative Diseases. *Dysphagia*. <https://doi.org/10.1007/s00455-020-10177-0>
- Dudik, J. M., Coyle, J. L., El-Jaroudi, A., Mao, Z.-H., Sun, M., & Sejdić, E. (2018). Deep learning for classification of normal swallows in adults. *Neurocomputing*, 285, 1-9. <https://doi.org/10.1016/j.neucom.2017.12.059>
- Dudik, J. M., Coyle, J. L., & Sejdic, E. (2015). Dysphagia Screening : Contributions of Cervical Auscultation Signals and Modern Signal-Processing Techniques. *IEEE Transactions on Human-Machine Systems*, 45(4), 465-477. <https://doi.org/10.1109/THMS.2015.2408615>
- Morinière, S., Boiron, M., Alison, D., Makris, P., & Beutter, P. (2008). Origin of the Sound Components During Pharyngeal Swallowing in Normal Subjects. *Dysphagia*, 23(3), 267-273. <https://doi.org/10.1007/s00455-007-9134-z>
- So, B. P.-H., Chan, T. T.-C., Liu, L., Yip, C. C.-K., Lim, H.-J., Lam, W.-K., Wong, D. W.-C., Cheung, D. S. K., Cheung, & J. C.-W. (2022). Swallow Detection with Acoustics and Accelerometric-Based Wearable Technology : A Scoping Review. *International Journal of Environmental Research and Public Health*, 20(1), 170. <https://doi.org/10.3390/ijerph20010170>



# Paramètres pertinents pour l'évaluation du niveau de lecture : vers l'élaboration d'un outil d'aide à la fluence

Laure DUCLOS<sup>1</sup>  
Elisa MABBOUX<sup>1</sup>  
Faraniaina RABARISOA<sup>3</sup>  
Muriel LALAIN<sup>1</sup>  
Corinne FREDOUILLE<sup>2</sup>  
Sophie HAMON<sup>3</sup>  
David RONDELLO<sup>3</sup>

<sup>1</sup> Aix-Marseille Univ, LPL, CNRS, Aix-en-Provence

<sup>2</sup> LIA, Avignon Université

<sup>3</sup> Société MobiDys, Nantes

Cette étude se donne pour objectif l'élaboration d'un outil numérique d'aide à la fluence en lecture.

Alors que la réussite de l'apprentissage de la lecture constitue un accomplissement intellectuel et un gage de réussite scolaire, son échec promet retards scolaires et difficultés d'intégration professionnelle et sociale. L'enjeu est tel que depuis 2021, le Conseil Scientifique de l'Education Nationale (CSEN) recommande d'évaluer les performances de lecture des élèves en début de 6<sup>ème</sup> au travers notamment d'une mesure de « fluence de lecture »<sup>11</sup>. Cette mesure, établie à partir d'un exercice de lecture à voix haute, rend compte de la vitesse et de la précision du décodage au travers d'un score égal au nombre de mots correctement lus en une minute.

Dans ce contexte où la fluence de lecture témoigne du niveau de lecture, la société MobiDys<sup>12</sup> s'est donné pour objectif d'intégrer aux livres numériques augmentés d'aide à la lecture qu'elle propose depuis 2016, un outil spécifique d'aide à la fluence. C'est autour de cet objectif que nous avons uni nos compétences pour 1) caractériser la fluence, 2) identifier les paramètres pertinents pour son évaluation et 3) intégrer les détections et mesures automatiques de ces paramètres dans les livres numériques.

L'acte de lire est défini comme la mise en œuvre de mécanismes spécifiques à la lecture (capacités phonologiques) et non spécifiques (capacités de compréhension) qui permettent de décoder et comprendre un texte lu (Gough & Tunmer, 1986). D'après Samuels (2006), la précision de lecture, la reconnaissance automatique des mots et l'expression orale (volume, intonation et rythme) témoignent de la compétence de lecture et caractérisent la lecture fluente. Ainsi, à l'instar du CSEN qui réduit la fluence aux paramètres de précision et de vitesse, plusieurs travaux ont montré l'importance de la prosodie et de la compréhension (Kuhn & Stahl, 2003, Schreiber, 1980, 1991, Rasinski, 2004). C'est ce que confirment les

---

<sup>11</sup> Note du CSEN, 20 mai 2021

<sup>12</sup> <https://www.mobidys.com/>

résultats préliminaires d'analyses d'un corpus de lecture de 61 élèves de 6<sup>ème</sup> âgés de 11 à 15 ans, normo-lecteurs (n=32) et TSA (n=29). Nous avons mesuré la précision et la vitesse de lecture, le nombre et la durée des pauses ainsi que les variations de fréquence fondamentale, au regard des scores de fluence établis selon les recommandations du CSEN. Les premiers résultats montrent une absence de corrélation entre la précision du décodage (fluence CSEN) et les variations de F0 (*Figure 1*) ce qui témoigne de l'importance de la prise en compte de ce paramètre prosodique pour évaluer le niveau de lecture.

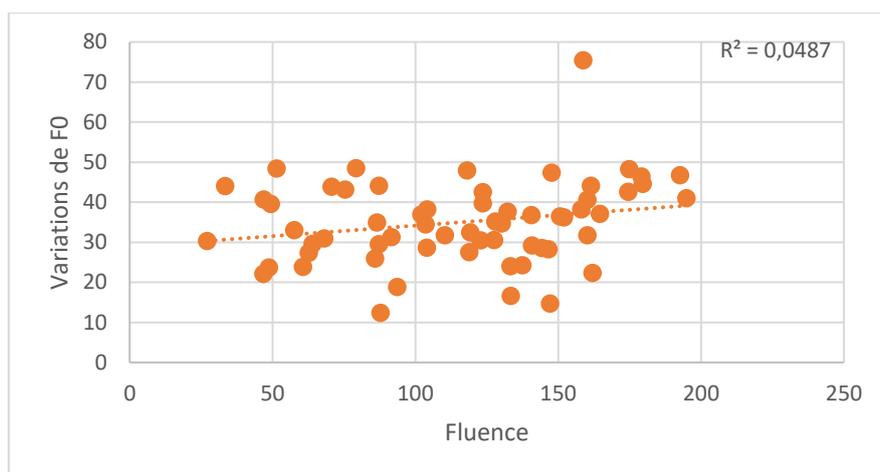


Figure 1. Relation entre le score de fluence et les variations de F0 en production de lecture

### Références bibliographiques

Gough, P.B., & Tunmer, W.E. (1986). Decoding, reading, and reading disability. *Remedial and Special Education*, 7, 6–10.

Kuhn, M.R., & Stahl, S.A. (2003). Fluency: A review of developmental and remedial practices. *Journal of Educational Psychology*, 95, 3–21.

Rasinski, T. V. (2004). *Assessing reading fluency*. Honolulu, HI: Pacific Resources for Education and Learning.

Samuels, S. J. (2006). Toward a model of reading fluency, In Samuels, S.J. & Farstrup, A.E. (Eds.), *What research has to say about fluency instruction*, Newark, DE: International Reading Association, 24-46.

Schreiber, P. A. (1980). On the acquisition of reading fluency. *Journal of Reading Behavior*, 12, 177-186.

Schreiber, P. A. (1991). Understanding prosody's role in reading acquisition. *Journal of Reading Behavior*, 12, 177-186.

# Comparaison de l'effet d'amorces rythmiques vocales, musicales et motrices sur le temps d'initiation de la parole chez des adultes tout-venant

Anne-Flore GINDRE<sup>1</sup>  
Halima SAHRAOUI<sup>1</sup>  
Corine ASTESANO<sup>1</sup>

<sup>1</sup>Laboratoire de NeuroPsychoLinguistique,  
EA4156,  
Université Toulouse 2 Jean Jaurès

L'altération du rythme de la parole est un indice de l'aphasie non fluente, la maladie de Parkinson ou encore du bégaiement et se traduit par une difficulté à initier la parole ou maintenir une parole fluide (Aichert et al., 2021 ; Brady, 1969). Certains protocoles utilisent ainsi l'entraînement rythmique pour la rééducation (Stahl et al., 2011 ; Thaut & Hoemberg, 2019). Dans les recherches étudiant l'impact du rythme sur la parole dans la population clinique et générale (Cason, Astésano & Schö, 2015a ; Cason et al., 2015b ; Gould et al., 2016 ; Norton et al., 2009 ; Roa et al., 2022), les différentes modalités rythmiques employées (langage, musique, battement de la main) ne sont généralement pas comparées. Afin de proposer des pistes de rééducation, il nous semble pertinent de vérifier d'abord auprès de participants tout-venant quelles sont les modalités rythmiques les plus efficaces. Cette étude a donc pour objectif d'observer l'effet de différentes conditions d'amorçage rythmique sur le temps d'initiation de la parole (**TI**).

Dix locuteurs francophones (24-34 ans, 5 femmes) ont lu à haute voix des phrases cibles précédées d'amorces rythmiques auditives (corpus partiellement tiré de Cason, Astésano & Schön, 2015a). L'amorçage varie selon l'Engagement Moteur du participant (**EM**) (frappes de l'index en rythme avec l'amorce entendue), le Type de l'Amorce (**TA**) et la Congruence Métrique entre l'amorce entendue et la phrase cible à lire (**CM**). Le *Tableau 1* synthétise les 12 conditions d'amorçage rythmique.

Tableau 1. Conditions d'amorçage rythmique

Engagement du sujet (EM)	Moteur	Type de l'Amorce (TA)	Congruence Métrique (CM)	
sans Moteur (-EM)	Engagement	vocal Langagier (L)	Congruent (C)	
			Incongruent (IC)	
		Non	vocal Langagier (NL)	Congruent (C)
				Incongruent (IC)
			Musical (percussions) (M)	Congruent (C)
				Incongruent (IC)
avec moteur (+EM)	engagement	vocal Langagier (L)	Congruent (C)	
			Incongruent (IC)	
		Non	vocal Langagier (NL)	Congruent (C)
				Incongruent (IC)
			Musical (percussions) (M)	Congruent (C)
				Incongruent (IC)

Nous utilisons pour notre analyse un modèle linéaire mixte (VD : TI ; Facteurs Fixés : EM, TA et CM ; Intercept Aléatoire : participants) avec le module Gamlj<sup>13</sup> de Jamovi (The Jamovi project, 2022). Les résultats préliminaires indiquent plusieurs tendances :

1. Effets simples :
  - a. **EM** a un effet significatif sur TI ( $F(1, 939)=11.365, p<0.001$ ), indiquant que la condition +EM raccourcit le TI par rapport à -EM (Figure 1).
  - b. **TA** a un effet significatif sur TI ( $F(2, 939)=7.43, p<0.001$ ), NL étant le plus facilitant par rapport à L et M (Figure 2).
  - c. **CM** n'a pas d'effet sur TI.
2. Effets d'interaction : les tests post-hoc entre **EM** et **TA** montrent que la direction de l'effet de EM est similaire pour tous les TA mais n'est significative que pour M ( $p=0.026$ ) (Figure 3).

L'utilisation d'un amorçage rythmique auditif pour faciliter l'initiation de la parole serait ainsi plus efficace lorsqu'il est accompagné d'un engagement moteur (+EM) (Figure 1), en particulier pour un amorçage musical (M) (Figure 3). Lorsque notre protocole sera complet (30 participants supplémentaires), nous nous attendons à retrouver cet effet facilitateur de l'engagement moteur également pour les autres types d'amorces, notamment NL, qui permet l'initiation de parole la plus rapide (Figure 2).

<sup>13</sup> General Analyses for Linear Modeles in Jamovi

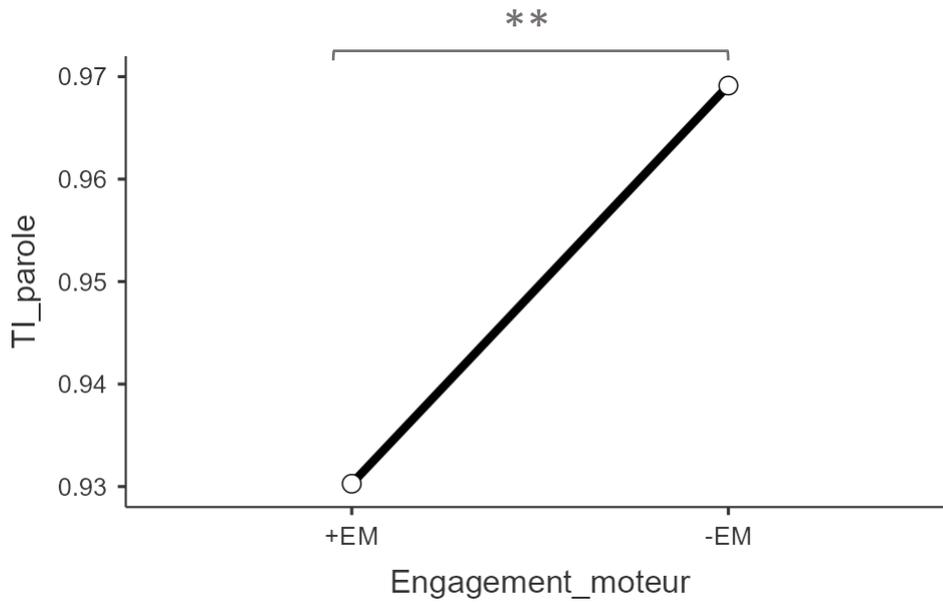


Figure 1. Comparaison du temps d'initiation de la parole (TI\_parole, en secondes), selon l'Engagement Moteur (Engagement\_moteur) du participant (+EM : avec Engagement Moteur, -EM : sans Engagement Moteur) ( $p < 0.001$ ) (valeurs modélisées).

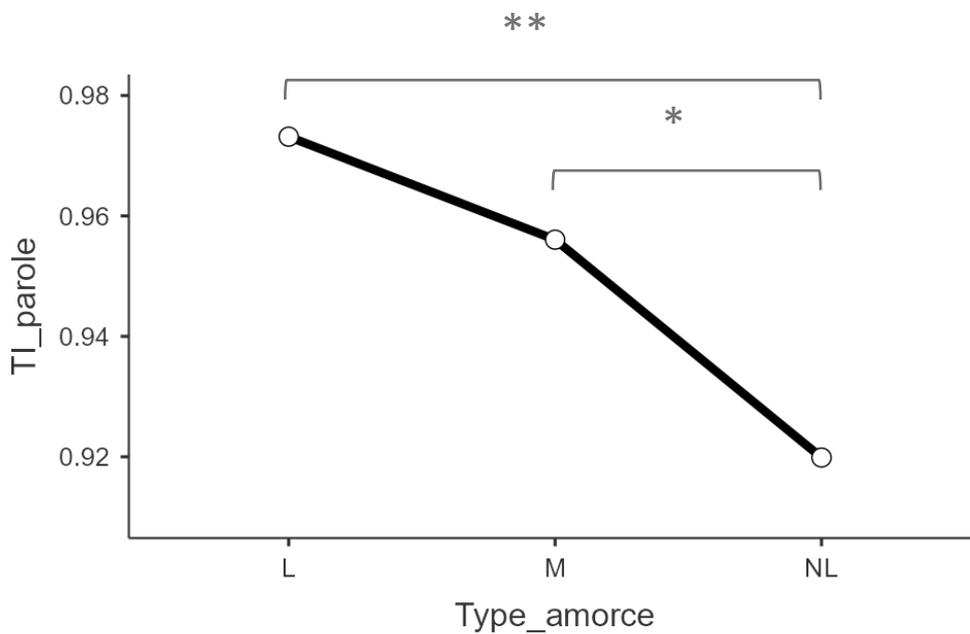


Figure 2. Comparaison du temps d'initiation de la parole (TI\_parole, en secondes), selon le Type d'Amorce (Type\_amorce) (L : vocal Langagier, NL : vocal Non Langagier, M : Musical) (L vs. NL :  $p < 0.001$  ; M vs. NL :  $p = 0.032$ ) (valeurs modélisées).

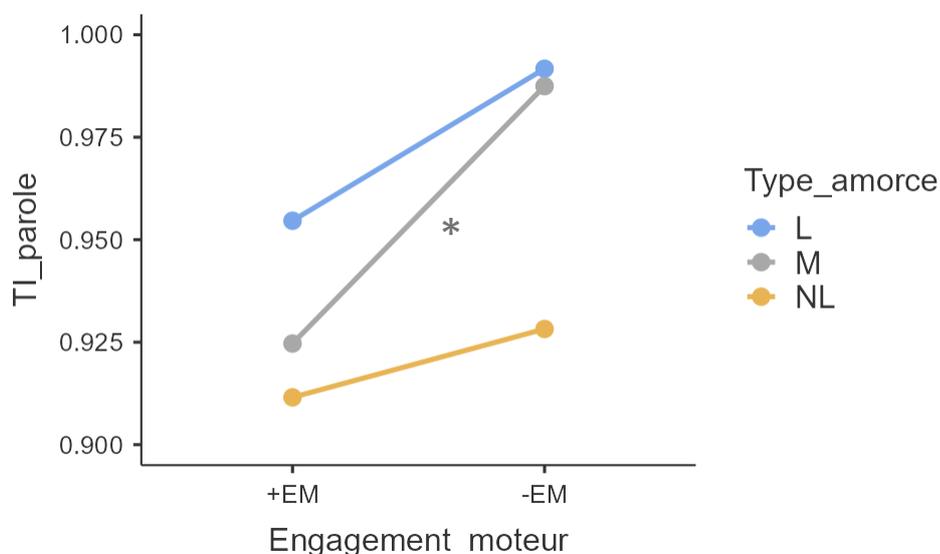


Figure 3. Comparaison du temps d'initiation de la parole (TI\_parole, en secondes), selon l'Engagement Moteur (Engagement\_moteur) du participant (+EM : avec Engagement Moteur, -EM : sans Engagement Moteur) et selon le Type d'Amorce (Type\_amorces) (L : vocal Langagier, NL : vocal Non Langagier, M : Musical). (M +EM vs. M -EM :  $p=0.026$ ) (valeurs modélisées).

### Références bibliographiques

- Aichert, I., Lehner, K., Falk, S., Späth, M., Franke, M., & Ziegler, W. (2021). In time with the beat: Entrainment in patients with phonological impairment, apraxia of speech, and parkinson's disease. *Brain Sciences*, 11(11). <https://doi.org/10.3390/brainsci11111524>
- Brady, J. P. (1969). Studies on the metronome effect on stuttering. *Behaviour Research and Therapy*, 7(2), 197–204. [https://doi.org/10.1016/0005-7967\(69\)90033-3](https://doi.org/10.1016/0005-7967(69)90033-3)
- Cason, N., Astésano, C., & Schön, D. (2015a). Bridging music and speech rhythm: Rhythmic priming and audio-motor training affect speech perception. *Acta Psychologica*, 155, 43-50. <https://doi.org/10.1016/j.actpsy.2014.12.002>
- Cason, N., Hidalgo, C., Isoard, F., Roman, S., & Schön, D. (2015b). Rhythmic priming enhances speech production abilities: Evidence from prelingually deaf children. *Neuropsychology*, 29(1), 102–107. <https://doi.org/10.1037/neu0000115>
- Gould, L., McKibben, T., Ekstrand, C., Lorentz, E., & Borowsky, R. (2016). The beat goes on: the effect of rhythm on reading aloud. *Language, Cognition and Neuroscience*, 31(2), 236–250. <https://doi.org/10.1080/23273798.2015.1089360>
- Norton, A. C., Zipse, L., Marchina, S., & Schlaug, G. (2009). Melodic Intonation Therapy: Shared Insights on How it is Done and Why it Might Help. *Stroke*, 1169, 431–436. <https://doi.org/10.1111/j.1749-6632.2009.04859.x>

- Roa, L. C., Mairano, P., Lacroix, J., Riegler, R., Moreau, C., & Basirat, A. (2022). *Does listening to non-linguistic rhythm impact speech production?* [communication par affiche]. 22nd conference of the European Society for Cognitive Psychology (ESCoP), Lille, France. hal-03763048
- Stahl, B., Kotz, S. A., Henseler, I., Turner, R., & Geyer, S. (2011). Rhythm in disguise: Why singing may not hold the key to recovery from aphasia. *Brain*, *134*(10), 3083–3093. <https://doi.org/10.1093/brain/awr240>
- Thaut, M. H., & Hoemberg, V. (2019). *Manuel clinique de rééducation par la musique*. De Boeck Supérieur.
- The jamovi project (2022). *jamovi*. (Version 2.3) [Computer software]. Retrieved from <https://www.jamovi.org>



## Session posters 2



# Glucides raffinés et attractivité vocale

Mélicca BARKAT-DEFRADAS<sup>1</sup>

Lulin YANG<sup>1</sup>

Michel RAYMOND<sup>1</sup>

Claire BERTICAT<sup>1</sup>

<sup>1</sup> Institut des Sciences de l'Évolution de Montpellier,  
CNRS, Université de Montpellier

Au cours du 20<sup>e</sup> siècle, un changement substantiel de régime alimentaire a eu lieu dans les pays occidentalisés avec, en particulier, l'introduction massive des glucides raffinés dont la consommation répétée a des conséquences délétères sur la santé (e.g. diabète II, obésité...). Les mécanismes physiologiques impliqués dans ces pathologies commencent à être bien connus et pourraient également affecter d'autres traits non examinés par la médecine, tels que les caractéristiques vocales. Afin d'investiguer cette question, nous avons évalué les variations de l'attractivité vocale perçue en fonction de la consommation de glucides raffinés des sujets enregistrés et des profils des juges-auditeurs.

Tout d'abord, 82 sujets (40 F et 42 H ; âge = 23 ans  $\pm$  1,5) ont été enregistrés en parole spontanée (i.e. description libre de leur habitat) après avoir renseigné un questionnaire relatif à différentes caractéristiques individuelles : âge, sexe, BMI, niveau socio-économique, habitudes alimentaires (permettant d'estimer la consommation de glucides raffinés hebdomadaire ainsi qu'un index de qualité alimentaire), activité physique et, pour les sujets de sexe féminin, prise éventuelle d'un contraceptif hormonal. Après mesure de différents paramètres acoustiques, les enregistrements de parole (durée moyenne = 7 secs  $\pm$  1,6) ont été manipulés à l'aide du logiciel Audacity© pour inverser le signal acoustique afin d'éviter que l'évaluation de l'attractivité vocale ne soit biaisée par l'information sémantique. Les stimuli ont ensuite été utilisés dans le cadre d'une expérience psychoacoustique. 133 juges de sexe opposé aux stimuli (74 H et 59 F, âge = 30 ans  $\pm$  13) ont évalué les échantillons vocaux présentés par paire de façon aléatoire par le biais d'une interface informatique (n= 1065 paires de voix féminines et 1466 paires de voix masculines). Ils/elles avaient pour tâche de choisir au sein de chaque paire de voix celle qu'ils/elles préféreraient. Les juges ont par ailleurs fourni différentes informations personnelles en lien avec leur profil sociologique (i.e., âge, sexe, niveau socio-éducatif, statut marital) et leurs habitudes alimentaires (cf. supra). La probabilité pour chaque voix d'être choisie a été analysée en fonction des caractéristiques des sujets via des modèles linéaires mixtes. Aucun effet des caractéristiques des juges n'a été détecté. Les résultats mettent en évidence un effet de la consommation chronique de glucides raffinés sur l'attractivité vocale, chez les femmes uniquement ( $p = .002$ ).



# Élaboration et normalisation d'une tâche évaluant les représentations sémantiques des noms et des verbes chez l'adulte

Pauline DEGHORAIN<sup>1</sup>

Lorraine CALIMEZ<sup>1</sup>

Anne-Sophie VAN DER MAREN<sup>2</sup>

Nathalie GUERRY<sup>2</sup>

Patrick FERY<sup>1,2</sup>

<sup>1</sup> Université Libre de Bruxelles, UR2NF

<sup>2</sup> HUB Hôpital Erasme, Service de Neuropsychologie et Logopédie

**Introduction** – Dans une approche amodale de la mémoire sémantique, un même système sémantique est contacté par les différentes modalités d'accès. Les représentations sémantiques sont organisées dans une approche en réseau et suivent une structure hiérarchique (Kumar, 2020). Ainsi, chaque concept est représenté comme un ensemble d'attributs interconnectés (Ursino et al., 2018). Deux grands types de propriétés sont mises en évidence : les attributs sensoriels ou physiques et les attributs fonctionnels et associatifs (Carbonnel, Charnalet & Moreaud, 2010). La littérature scientifique montre que la catégorie grammaticale (distinction substantif/verbe) (Vigliocco et al., 2010) et la catégorie sémantique (distinction biologique/manufacturée) (Monetta et al., 2020) affectent le traitement sémantique.

**Objectif** – Cette étude a pour objectif la création et la normalisation d'une nouvelle batterie pour évaluer précisément les capacités qui traduisent l'intégrité des représentations sémantiques chez des sujets sains de 50 à 65 ans.

**Méthode** – Le test élaboré consiste en une tâche d'appariement sémantique informatisée entre des items et leurs attributs. Les 105 items utilisés proviennent des mémoires de Ricca et Simonin (2019) et Ducroux (2020). La sélection de ces items repose sur un contrôle de la fréquence lexicale, de la catégorie grammaticale et de la catégorie sémantique. Le test a été présenté sous deux modalités : la modalité visuelle picturale et la modalité visuelle écrite. La tâche d'appariement vise à évaluer les attributs perceptifs et les attributs fonctionnels/associatifs de substantifs mais aussi de verbes. Les normes ont été récoltées sur un échantillon de 54 sujets sains et les variables démographiques du genre et de la scolarité ont été contrôlées.

**Résultats** – L'analyse des résultats montre que la tâche a un haut taux de réussite. En effet, 93 % des items utilisés sont réussis par au moins la moitié de l'échantillon. L'analyse des variables démographiques n'a pas révélé d'effet du genre significatif, mais a indiqué un effet du niveau d'études sur certains items. L'analyse des différents types d'items a souligné un effet principal de la modalité de présentation (écrite > picturale), de la catégorie grammaticale (noms > verbes), de la catégorie sémantique (manufacturé > biologique) et du type d'attribut (fonctionnels/associatif > perceptifs) sur le taux de réussite et les temps de réponse.

Conclusion – Ce travail de recherche a permis de recueillir des normes complètes sur l'intégrité des représentations sémantiques de noms mais aussi de verbes chez des adultes sains. Ces dernières devront être affinées et un recueil devra être réalisé sur une population pathologique comparable afin de déterminer la validité de la tâche créée.

### ***Références bibliographiques***

- Carbonnel, S., Charnallet, A., & Moreaud, O. (2010). Organisation des connaissances sémantiques : des modèles classiques aux modèles non abstraits. *Revue de neuropsychologie*, 2(1), 22-30. doi:10.3917/rne.021.0022
- Ducroux, M. (2020). Création et normalisation d'une batterie d'évaluation des systèmes mnésiques pré-sémantiques chez l'adulte (Mémoire). Université Libre de Bruxelles.
- Kumar, A. A. (2021). Semantic memory: A review of methods, models, and current challenges. *Psychon Bull Rev*, 28(1), 40-80. doi:10.3758/s13423-020-01792-x
- Monetta, L., Légaré, A., Macoir, J., & Wilson, M. (2020). Questionnaire Sémantique de Québec (QueSQ). Développement, validation et normalisation. *Canadian Journal on Aging / La Revue Canadienne Du Vieillessement*, 39(1), 98-106. doi:10.1017/S0714980819000333
- Ricca, A., & Simonin, P. (2019). Élaboration et normalisation d'une tâche évaluant le système sémantique chez l'adulte (Mémoire). Université Libre de Bruxelles.
- Ursino, M., Cuppini, C., Cappa, S. F., & Catricalà, E. (2018). A feature-based neurocomputational model of semantic memory. *Cogn Neurodyn*, 12(6), 525-547. doi:10.1007/s11571-018-9494-0
- Vigliocco, G., Vinson, D., Druks, J., Barber, H., & Cappa, S. (2010). Nouns and verbs in the brain: A review of behavioural, electrophysiological, neuropsychological and imaging studies. *Neuroscience and biobehavioral reviews*, 35, 407-426. doi:10.1016/j.neubiorev.2010.04.007

# La somnolence influence-t-elle la production des schwas chez les patients hypersomniaques ?

Colleen BEAUMARD<sup>1,2</sup>

Vincent P. MARTIN<sup>1,2</sup>

Yaru WU<sup>3,4,5</sup>

Jean-Luc ROUAS<sup>1</sup>

Pierre PHILIP<sup>2</sup>

<sup>1</sup>LaBRI, UMR 5800 CNRS, Bordeaux INP, Université de Bordeaux

<sup>2</sup>SANPSY, UMR 6033 CNRS, Université de Bordeaux

<sup>3</sup>CRISCO/EA4255, Université de Caen

<sup>4</sup>Laboratoire de Phonétique et de Phonologie, UMR 7018 CNRS, Université Sorbonne Nouvelle

<sup>5</sup>LISN, UMR 9015 CNRS, Université de Paris Saclay

Les cliniciens ont besoin de surveiller et évaluer les traitements des patients hypersomniaques en conditions écologiques. Le schwa étant optionnel (Durand, 2014 ; Burki et al., 2011) et pouvant être utilisé comme une hésitation en français, nous avons voulu déterminer si le niveau de somnolence des patients hypersomniaques influençait le nombre de schwas prononcés.

Nous avons utilisé un sous-corpus de 20 patients du corpus TILE (Martin et al., 2020) contenant des enregistrements de lecture de texte à voix haute collectés avant chaque opportunité de sieste d'un Test Itératif de Latence d'Endormissement (TILE) (Arand et al., 2005) au Service Universitaire de Médecine du Sommeil au CHU de Bordeaux. Il y a 5 opportunités de sieste lors du TILE, notre sous-corpus contient donc 100 enregistrements. Deux dimensions de somnolence, subjective et physiologique (Akerstedt & Gillberg, 1990), ont été mesurées à court- (lectures) et long-terme (locuteurs).

L'annotation des schwas a été réalisée manuellement sur les enregistrements en se référant aux transcriptions phonologiques du Lexique 3.83 (New et al., 2004). Nous avons calculé six descripteurs rapportés dans le *Tableau 1* avec leur définition, leur moyenne, et leur écart-type.

Tableau 1. Définition et moyenne (écart-type) des six descripteurs.

<i>Descripteur</i>	<i>Définition</i>	<i>Moy (e.t.)</i>
%S	Ratio du nombre de schwas prononcés en fonction du nombre théorique de schwas	88,0 (9,8)
%S2C	Ratio du nombre de schwas prononcés dans les mots avec 2 consonnes phonétiques ou plus en fonction du nombre théorique de schwas	27,2 (9,5)
%S1C	Ratio du nombre de schwas prononcés dans les mots avec 1 consonne phonétique ou moins en fonction du nombre théorique de schwas	60,7 (15,6)
%Sa	Ratio du nombre de schwas prononcés (non présents) en fonction du nombre théorique de schwas	7,8 (10,9)
%S2Ca	Ratio du nombre de schwas prononcés (non présents) dans les mots avec 2 consonnes phonétiques ou plus en fonction du nombre théorique de schwas	60,3 (41,8)
%S1Ca	Ratio du nombre de schwas prononcés (non présents) dans les mots avec 1 consonne phonétique ou moins en fonction du nombre théorique de schwas	19,7 (30,4)

Afin d'identifier les facteurs influençant majoritairement la prononciation des schwas dans la lecture à voix haute, une ANOVA multivariée à mesures répétées a été réalisée (Tableau 2). Les corrélations univariées avec un effet significatif entre les descripteurs et les facteurs y sont présentées.

Tableau 2. Haut : résultats de l'ANOVA multivariée à mesures répétées. Seules les variations intra-locuteur sont reportées, aucun effet significatif n'ayant été observé dans les variations inter-locuteur. Bas : corrélation entre les moyennes de chaque descripteur en fonction des locuteurs et des caractéristiques des locuteurs.

td. :  $p < .1$  ; \* :  $p < .05$  ; \*\* :  $p < .01$  ; \*\*\* :  $p < .001$

<i>Facteur</i>	<i>%S</i>	<i>%S2C</i>	<i>%S1C</i>	<i>%Sa</i>	<i>%S2Ca</i>	<i>%S1Ca</i>
Somnolence subjective à court-terme	*	***	***			*
Somnolence physiologique à court-terme		***	*			
Textes	***	***	***	***	***	
Somnolence subjective à court-terme : Textes	td.					
<i>Corrélations : <math>\rho</math> (sig.) de Spearman</i>						
IMC				.50 (*)		
Somnolence subjective à long-terme			-.45 (*)			

Des effets significatifs ont été observés entre les variations intra-locuteur de somnolence subjective et physiologique et celles des descripteurs (Figures 1 et 2). Les descripteurs des schwas présents prononcés (excepté %S1Ca) semblent plus liés à la somnolence subjective et physiologique que les descripteurs des schwas ajoutés.

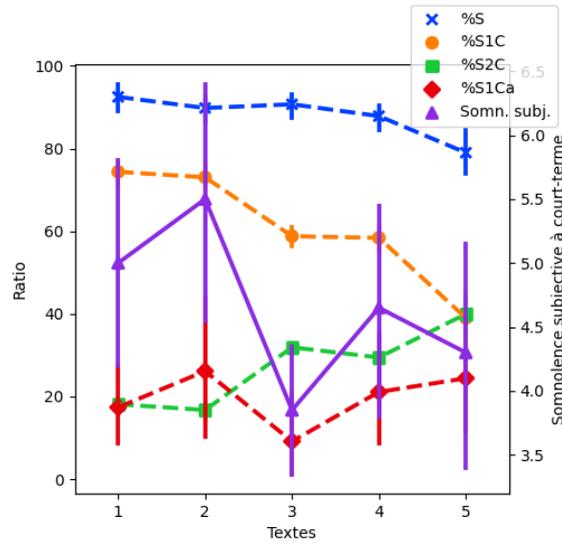


Figure 1. Variations intra-locuteur de la somnolence subjective à court-terme et des descripteurs ayant un effet significatif d'après l'ANOVA.

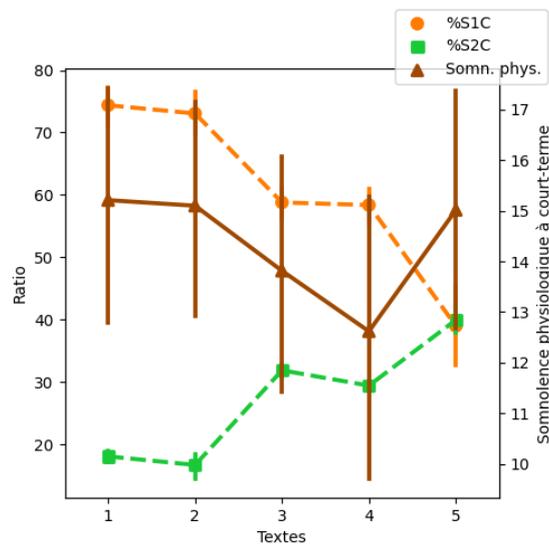


Figure 2 Variations intra-locuteur de la somnolence physiologique à court-terme et des descripteurs ayant un effet significatif d'après l'ANOVA.

En plus de l'ANOVA, deux classifications (*Tableaux 3 et 4*) en validation croisée *Leave-One-Speaker-Out (LOSO)* pour la détection de la somnolence à court- et long-terme ont été réalisées avec tous les descripteurs. Deux techniques de fusion de descripteurs ont été testées : moyennés et agrégés. La mesure « Somnolence physiologique à long-terme » n'ayant aucun patient considéré comme somnolent, aucun résultat n'a pu être obtenu pour cette mesure. Les résultats vont dans le sens d'un lien entre la prononciation des schwas et la somnolence subjective.

Tableau 3. Résultats de la classification avec une régression logistique pour la détection de la somnolence à court-terme sur les 100 enregistrements (LOSO). UAR : *Unweighted Average Recall*

<i>Référence</i>	<i>Mesure</i>	<i>#somnolent</i>	<i>UAR%</i>
A	Somnolence subjective à court-terme	33	<b>61,5</b>
B	Somnolence physiologique à court-terme	20	50,0

Tableau 4. Résultats de la classification avec une régression logistique pour la détection de la somnolence à long-terme sur les 20 locuteurs (LOSO). Mo: Moyennés; Ag: Agrégés ; UAR : *Unweighted Average Recall*

<i>Référence</i>	<i>Mesure</i>	<i>#somnolent</i>	<i>Mo/Ag</i>	<i>UAR%</i>
1a	Somnolence subjective à court-terme moyennée	6	Mo	<b>63,1</b>
1b			Ag	<b>63,1</b>
2a	Somnolence subjective à long-terme	10	Mo	55,0
2b			Ag	<b>65,0</b>

Pour conclure, il semble y avoir un lien entre production des schwas et la somnolence subjective et physiologique lors de lecture à voix haute. Ces résultats nous encouragent à automatiser le processus avec la Reconnaissance Automatique de la Parole afin de détecter les schwas avant de l'appliquer à la parole spontanée.

### Références bibliographiques

- Akerstedt, T., & Gillberg, M. (1990). Subjective and objective sleepiness in the active individual. *International Journal of Neuroscience*.
- Arand, D., Bonnet, M., Hurwitz, T., Mitler, M., Rosa, R., & Sangal, R. B. (2005). The clinical use of the MSLT and MWT. *Sleep*.
- Burki, A., Ernestus, M., Gendrot, C., Fougéron, C., & Frauenfelder, U. H. (2011). What affects the presence versus absence of schwa and its duration: A corpus analysis of French connected speech. *The Journal of the Acoustical Society of America*.
- Durand, J. (2014). A la recherche du schwa : données, méthodes et théories. SHS Web of Conferences.
- Martin, V. P., Rouas, J.-L., Micoulaud-Franchi, J.-A. & Philip, P. (2020). The Objective and Subjective Sleepiness Voice Corpora. Marseille: LREC.
- New, B., Pallier, C., Brysbaert, M., & Ferrand, L. (2004). Lexique 2: A new French lexical database. *Behavior Research Methods, Instruments, & Computers*.

# Validation d'un corpus de paires minimales illustrées adapté aux enfants de 7 à 10 ans : Paires minimales Illustrées Francophones (PIF)

Amelia PETTIROSSI<sup>1,2</sup>  
Nicolas AUDIBERT<sup>1</sup>  
Lise CREVIER-BUCHMAN<sup>1,3</sup>

<sup>1</sup> Laboratoire de Phonétique et Phonologie,  
UMR7018 Université Sorbonne Nouvelle/CNRS, Paris

<sup>2</sup> Imagerie & Cerveau,  
UMR1253 INSERM/Université de Tours, Tours

<sup>3</sup> Hôpital Foch : Service de Laryngologie Phoniatrie,  
Université Paris Saclay, Suresnes

Dans le cadre d'entretiens non-structurés des planches de bande-dessinées sans dialogue tel que « La chute dans la boue » ou « Frog, where are you ? » sont respectivement utilisées à des fins de bilans orthophoniques (Chevrie-Muller & Plaza, 2001) ou d'études cliniques (Reilly, 2004 ; Ash et al., 2006). Dans cette étude nous proposons de mettre à disposition de la communauté une banque de données d'images testée sur des enfants et pouvant être utilisées dans différentes tâches expérimentales afin de pallier le manque d'illustrations libres de droits exploitables.

Une phase de sélection a été réalisée à partir d'un lexique cible de 33 noms communs français de structure syllabique CVC et représentés à travers 37 dessins. Trois enfants de 7 ans sont interrogés dans un centre culturel du Val-de-Marne (94). Premièrement, les enfants participent à une tâche de dénomination d'images dans laquelle 14 mots sont produits selon la cible par tous (*Tableau 1*). Deuxièmement, une tâche d'identification d'image est réalisée, nous observons alors que 10 des mots-cibles restant ne sont pas connus par tous les enfants alors que 13 autres sont parfaitement identifiés (*Tableau 2*).

Les 27 images correctement dénommées ou identifiées par tous les enfants constituent la banque d'image PIF (Paires minimales Illustrées Francophones) mise à disposition sous licence Creative Commons CC-BY-NC-SA au lien suivant : <https://hal.science/hal-03945324/>.

Parmi ces 27 items, cinq paires minimales associées à leur dessin respectif ([bak]/[pak]/, [bul]/[mul] (*gâteau*), [muʃ]/[mus], [pil]/[pyl] et [pus]/[pys]) et produites par des locutrices dans la phrase porteuse « Clique sur le dessin de... » ont déjà été utilisées dans un test de discrimination à destination de 22 enfants de 7 à 10 ans (Pettirossi, 2021). Les résultats suivants se concentrent sur les productions de 10 locutrices sans trouble vocal, évaluées G0 sur l'échelle GRBAS (Hirano, 1981). Chaque mot-cible étant présenté 20 fois, nous basons nos analyses sur 4400 réponses au total. Nous n'observons que 8.36% d'erreur de discrimination, avec un effet significatif de la paire minimale testée ( $F(4)=14.46$  ;  $p<.0001$ ).

La paire [pak]/[bak], représentant 33% des erreurs, est moins bien reconnue que les autres paires ( $p=0.01$ ). Nous trouvons également une différence significative entre les tranches d'âge ( $F(3)=6.07$  ;  $p=0.0004$ ), car les enfants de 7 et 8 ans font significativement plus d'erreurs d'identification que ceux de 9 et 10 ans.

Ces résultats confirment que les dessins proposés permettent aux enfants de 7 à 10 ans d'identifier les mots-cibles correspondants.

Tableau 1. Nombre d'enfants ayant correctement produit le mot-cible dans une tâche de dénomination d'image (*Italique* : Deux dessins sont testés pour le même mot)

Illustration désignée par le mot-cible par 3 enfants sur 3	Illustration désignée par le mot-cible par 2 enfants sur 3	Illustration désignée par le mot-cible par 1 enfant sur 3	Illustration désignée par le mot-cible par 0 enfant sur 3
Bottes [bɔt], Bouche [buʃ], Boule [bul], Bulle [byl], Bus [bys], Mare [maʁ], Mille [mil], <i>Mousse (bain) [mus]</i> , Pâques [pak], Patte [pat], Pouce [pus], Poule [pul], Puce [pus], Pull [pyl]	Bac (à sable) [bak], Barre [baʁ], Mouche [muʃ], <i>Moule (gâteau) [mul]</i> , Pile [pil]	<i>Mèche (coupe) [mɛʃ]</i> , <i>Mèche (colorées) [mɛʃ]</i>	Balle [bal], Batte [bat], Mage [maʒ], Malle [mal], Maths [mat], Mère [mɛʁ], Motte [mɔt], <i>Moule (mollusque) [mul]</i> , <i>Mousse (savon) [mus]</i> , Mule [myl], Page [paʒ], Pale [pal], <i>Pêche (fruit) [pɛʃ]</i> , <i>Pêche (poisson) [pɛʃ]</i> , Père [pɛʁ], Potes [pɔt]

Tableau 2. Illustrations correctement reconnues, ou non, lors d'une tâche d'identification d'image

Illustration correctement identifiée grâce au mot-cible par tous les enfants	Illustration mal identifiée avec présentation du mot-cible par au moins 1 enfant
Bac (à sable) [bak], Balle [bal], Maths [mat], <i>Mèche (colorées) [mɛʃ]</i> , Mère [mɛʁ], Mouche [muʃ], <i>Moule (gâteau) [mul]</i> , <i>Moule (mollusque) [mul]</i> , Page [paʒ], <i>Pêche (poisson) [pɛʃ]</i> , <i>Pêche (fruit) [pɛʃ]</i> , Père [pɛʁ], Pile [pil]	Batte [bat], Barre [baʁ], Mage [maʒ], Malle [mal], <i>Mèche (coupe) [mɛʃ]</i> , <i>Mousse (savon) [mus]</i> , Motte [mɔt], Mûle [myl], Pale [pal], Potes [pɔt]

## Références bibliographiques

- Ash, S., Moore, P. A., Antani, S., McCawley, G., M, W., & Grossman, M. (2006). Trying to tell a tale : Discourse impairments in progressive aphasia and frontotemporal dementia. *Neurology*, *66*(9), 1405-1413. <https://doi.org/10.1212/01.wnl.0000210435.72614.38>
- Chevrié-Muller, C. & Plaza, M. (2001). *N-EEL: nouvelles épreuves pour l'examen du langage*, ECPA, Les Editions du Centre de psychologie appliquée.
- Hirano, M. (1981). Clinical examination of voice, *Disorders of human communication*, *5*, 1-99.
- Pettirossi, A. (2021). La dysphonie chez les professeures des écoles: perception et représentations, Thèse de doctorat. Université de la Sorbonne nouvelle-Paris III.
- Reilly, J. (2004). “Frog, where are you ? ” Narratives in children with specific language impairment, early focal brain injury, and Williams syndrome. *Brain and Language*, *88*(2), 229-247. [https://doi.org/10.1016/s0093-934x\(03\)00101-9](https://doi.org/10.1016/s0093-934x(03)00101-9)



# Perception de parole chez les enfants porteurs d'implants cochléaires : Apports de l'Auditory Verbal Therapy et de la Langue française Parlée Complétée

Lucie VAN BOGAERT<sup>1,2</sup>  
Laura MACHART<sup>1,2</sup>  
Hélène LÈVENBRUCK<sup>1</sup>  
Anne VILAIN<sup>2</sup>

<sup>1</sup> Laboratoire de Psychologie et NeuroCognition,  
UMR 5105 CNRS,  
Univ. Grenoble Alpes,  
Univ. Savoie Mont-Blanc

<sup>2</sup> GIPSA-Lab,  
UMR 5216 CNRS,  
Grenoble INP,  
Univ. Grenoble Alpes

Bien que les implants cochléaires (IC) améliorent la perception de la parole, les informations auditives transmises par celui-ci restent dégradées et imparfaites (Leybaert & LaSasso, 2010). Cette étude porte sur les bénéfices pour la perception de parole de deux approches de soutien de la langue vocale chez les enfants porteurs d'IC. L'Auditory Verbal Therapy (AVT) est un programme d'intervention précoce qui s'appuie principalement sur l'entraînement auditif pour que l'enfant sourd puisse développer des compétences auditives et langagières adéquates. La Langue française Parlée Complétée (LfPC) est un outil d'aide à la communication qui utilise la multisensorialité de la parole pour désambigüiser la lecture labiale en y ajoutant un geste manuel.

Cette étude inclut des enfants âgés de 5 à 11 ans : 90 enfants normo-entendants (NE), 9 enfants porteurs d'IC ayant participé à un programme AVT (AVT), 6 enfants porteurs d'IC ayant un niveau élevé de décodage de la LfPC (LfPC+), et 19 enfants porteurs d'IC ayant un faible niveau de décodage de la LfPC (LfPC-). Le niveau de décodage a été mesuré en utilisant le test TERMO (Busquet & Descourtieux, 2000). Dans cette étude, la tâche de jugement de lexicalité de la batterie EULALIES (Meloni et al. 2017) a été utilisée. Des mots étaient présentés en modalité audiovisuelle, avec altération phonologique ou non et l'enfant devait juger si ceux-ci étaient prononcés correctement. La sensibilité de la perception a été mesurée par un score *d'* (Macmillan & Creelman, 2004). Un intervalle de prédiction à 95% a été calculé pour les enfants NE et les scores des enfants porteurs d'IC ont été comparés à cet intervalle (par des tests de conformité des proportions). Pour les scores IC en dehors de cet intervalle, la distance par rapport à la borne supérieure de l'intervalle a été calculée.

Les résultats montrent que les enfants LfPC+ et LfPC- ont des performances significativement plus faibles que leurs pairs entendants et que les enfants AVT ont tendance à avoir des scores plus faibles. De plus, les scores des enfants des groupes AVT et LfPC+

sont plus proches de l'intervalle que ceux des enfants du groupe LfPC-. Ainsi, cette étude semble montrer que ces deux approches (AVT et LfPC) contribuent à la rééducation de la perception de parole, et souligne l'importance d'utiliser une approche spécifique en soutien de PIC.

### Références bibliographiques

- Busquet, D., Descourtieux, C. (2000). TERMO Test d'évaluation de la réception du message oral par l'enfant sourd. Unithèque. <https://www.unitheque.com/termo-test-evaluation-reception-message-oral-par-enfant-sourd/ortho/Livre/4074>
- Leybaert, J., LaSasso, C.J. (2010). Cued Speech for Enhancing Speech Perception and First Language Development of Children With Cochlear Implants. *Trends Amplif*, 96–112.
- MacMillan, N.A., Creelman, C.D. (2004). Detection Theory: A User's Guide, 2nd ed. *Psychology Press*, New York.
- Meloni, G., Loevenbruck, H., Vilain, A. MacLeod A. (2017). EULALIES, The France-Québec Speech Sound Disorders project, Actes de 14th International Congress for the Study of Child Language (IASCL), Lyon, France.

# La voix comme marqueur potentiel de l'autisme

Frédéric BRIEND<sup>1,2\*</sup>  
Céline DAVID<sup>2</sup>  
Silvia SILLERESI<sup>3</sup>  
Joëlle MALVY<sup>1,2</sup>  
Sandrine FERRÉ<sup>1,2</sup>  
Marianne LATINUS<sup>1,2,4\*</sup>

<sup>1</sup> EXAC·T, Centre Universitaire de Pédopsychiatrie, CHRU de Tours,  
Tours, France

<sup>2</sup> UMR 1253, iBrain, Université de Tours, INSERM, 37000 Tours, France

<sup>3</sup> University of Milano-Bicocca, Department of Psychology, Milan, Italy

<sup>4</sup> Centro de Estudios en Neurociencia Humana y Neuropsicología.  
Facultad de Psicología. Universidad Diego Portales,  
Santiago, Chile

Le dépistage précoce d'enfant avec un trouble du spectre de l'autisme (TSA) est crucial pour un diagnostic précis et une intervention spécifique, permettant des effets positifs à long terme sur leurs symptômes et habiletés. Le besoin d'outils diagnostiques objectifs est particulièrement important du fait du faible pouvoir diagnostique de ceux existants actuellement. Une prosodie atypique, évaluée par des outils diagnostiques tels que l'ADOS, est l'un des premiers marqueurs de l'autisme. Un cri atypique à la naissance est souvent associé à un dysfonctionnement du système nerveux central (LaGasse et al., 2005). Des études précédentes, reposant sur une approche univariée, semblent suggérer des différences acoustiques entre les voix des personnes autistes et des personnes ayant un développement typique ; les résultats restent cependant contradictoires (Fusaroli et al., 2017). Dans cette étude, nous cherchions à évaluer si les paramètres acoustiques de la voix pouvaient permettre de discriminer des enfants avec autisme d'une population hétérogène d'enfants : enfants au développement typique (DT), enfants sourds avec implants cochléaires (SIC) et enfants avec un trouble du développement du langage (TDL). Cent-huit enfants (entre 6 et 10 ans), dont 38 enfants autistes, 24 enfants typiques et 46 enfants présentant un autre trouble (TDL et SIC) ont participé à notre étude. Les propriétés acoustiques d'échantillons de paroles recueillis dans le cadre d'une tâche de répétition de non-mots ont été mesurées (Dos Santos & Ferre, 2018). Nous avons utilisé une analyse en cluster de type k-means, que nous avons supervisée par une analyse ROC (Nikas & Low, 2011), associée à une validation croisée de Monte-Carlo, pour développer un modèle de classification qui pourrait discriminer un enfant de diagnostic inconnu.

Nous montrons que les paramètres acoustiques de la voix permettent de classer les enfants autistes avec une performance avoisinant les 91% en comparaison à un groupe homogène d'enfant ayant un développement typique, et autour de 85% en comparaison à une population hétérogène d'enfants non-autistes (DT, TDL, SIC). Les paramètres

acoustiques discriminant sont le shimmer moyen et la fréquence du premier formant. Les performances de l'algorithme k-means supervisé associé à une validation croisée de Monte Carlo sont supérieures à celles rapportées précédemment. Nos résultats montrent que des paramètres acoustiques, faciles à mesurer, pourraient être utilisés comme outils d'aide au diagnostic, spécifique de l'autisme.

### Références bibliographiques

- Dos Santos, C., & Ferre, S. (2018). A Nonword Repetition Task to Assess Bilingual Children's Phonology. *Language Acquisition*, 25, 58-71.
- Fusaroli, R., Lambrechts, A., Bang, D., Bowler, D. M., & Gaigg, S. B. (2017). "Is voice a marker for Autism spectrum disorder? A systematic review and meta-analysis". *Autism Res*, 10(3), 384-407. <https://doi.org/10.1002/aur.1678>
- LaGasse, L. L., Neal, A. R., & Lester, B. M. (2005). Assessment of infant cry: acoustic cry analysis and parental perception. *Ment Retard Dev Disabil Res Rev*, 11(1), 83-93. <https://doi.org/10.1002/mrdd.20050>
- Nikas, J. B., & Low, W. C. (2011). Application of clustering analyses to the diagnosis of Huntington disease in mice and other diseases with well-defined group boundaries. *Comput Methods Programs Biomed*, 104(3), e133-147. <https://doi.org/10.1016/j.cmpb.2011.03.004>

# Les patients en état de coma sont-ils sensibles à la prosodie ?

Estelle PRUVOST-ROBIEUX<sup>1,2,3</sup>  
Coralie JOUCLA<sup>4</sup>  
Rudradeep GUHA<sup>4</sup>  
Sarah BENGHANEM<sup>1,2,5</sup>  
Martine GAVARET<sup>1,2,3</sup>  
Jean-Julien AUCOUTURIER<sup>4</sup>

<sup>1</sup> Institut Paris Neurosciences et Psychiatrie (IPNP), INSERM, Paris, France

<sup>2</sup> Université de Paris Cité, Paris, France

<sup>3</sup> Service de neurophysiologie clinique et épileptologie,  
GHU Paris Psychiatrie et Neurosciences, Paris, France

<sup>4</sup> Université de Franche-Comté, SUPMICROTECH, CNRS,  
Institut FEMTO-ST, Besançon, France

<sup>5</sup> Service de médecine intensive et réanimation, Hôpital Cochin, APHP,  
Paris, France

L'évaluation du niveau de conscience des patients en état de coma comprend des paradigmes auditifs de type *oddball*, mêlant le prénom du patient (utilisé comme cible) et de fréquentes bouffées tonales (non cibles) (André-Obadia et al., 2018). Ces paradigmes élicitent une réponse centro-pariétale P300, en faveur d'un bon pronostic d'éveil. Actuellement toutefois les caractéristiques acoustiques des prénoms utilisés ne sont pas contrôlées et la sensibilité à la prosodie des patients comateux n'a pas été démontrée.

Dans une étude rétrospective (Pruvost-Robieux et al., 2022), nous avons analysé les enregistrements sonores et les P300 résultantes de 251 patients comateux à Lyon (Hospices civils) et Paris (GHU Paris, Sainte Anne). Une analyse de *reverse-correlation* entre la prosodie du prénom, a posteriori, et les caractéristiques de la P300 a montré que les prénoms prononcés avec une prosodie ascendante (versus descendante) étaient associés à une latence de réponse P300 plus courte (- 66 ms; IC95%: [6.36 - 125.9ms]).

Suite à ces résultats, nous avons mené une étude prospective chez 24 volontaires sains, évaluant l'impact de variants prosodiques du prénom sur les potentiels évoqués auditifs tardifs dans 2 paradigmes *oddball* successifs : un ciblant des variantes de fréquence fondamentale (prénom avec prosodie ascendante, descendante ou plate), l'autre des variantes de timbre (prénom prononcé par une voix souriante, rugueuse ou neutre). Ces variants d'expressivité du prénom ont été générés de façon standardisée par un algorithme (Arias et al., 2021), à partir d'un seul enregistrement neutre. Durant le paradigme, il était demandé aux participants de compter mentalement le nombre de prénoms entendus indépendamment de leur prosodie. Un enregistrement électro-encéphalographique (EEG) était réalisé concomitamment.

Dans cette étude, contrairement à l'étude rétrospective, la prosodie ascendante du prénom était cette fois associée à des réponses de détection du prénom significativement *moins* amples sur les dérivations pariétales que les prosodies descendantes (sur Pz, amplitude = 1.6 vs 2.4 $\mu$ V,  $p=0.02$ ) et plates. Nous n'avons par ailleurs pas mis en évidence de différence significative entre aucune variante de timbre du prénom.

Notre interprétation est que les deux études reflètent un seul et même effet exécutif de capture attentionnelle par la prosodie ascendante des prénoms, traitée comme plus alertante. Dans la tâche passive, les prénoms alertants potentialisent la réponse P300 ; dans la tâche active, ceux-ci interfèrent avec l'attention nécessaire au comptage mental, tâche pour laquelle la prosodie du prénom doit être ignorée.

Ces deux études mettent donc en évidence une influence de la prosodie sur les caractéristiques des réponses P300 enregistrées en EEG. La prosodie doit être contrôlée dans les paradigmes cognitifs acoustiques. Une étude prospective est actuellement en cours sur l'influence de la prosodie du prénom chez les patients comateux.

**Figure 1 : La prosodie ascendante du prénom est associée à une réponse P300 moins ample sur Pz.**

- a) Amplitude des réponses obtenues sur Pz sur la fenêtre [-100; +600] ms après la stimulation prénom, dans les conditions prosodie ascendante (vert) ou descendante (rouge). Une analyse de permutation de cluster met en évidence une différence significative entre les courbes obtenues sur la fenêtre de temps [180;320 ms].
- b) Comparaison des amplitudes en Pz sur la fenêtre de temps [180;320 ms]. L'amplitude obtenue pour les prénoms de prosodie ascendante est significativement moindre que celle des prénoms de prosodie descendante (1,6 versus 2,4  $\mu$ V respectivement;  $p = 0,02$ ).
- c) Cartographie d'amplitude des réponses enregistrées sur la fenêtre [180-320 ms] dans les conditions prosodie descendante et ascendante et analyse de permutation de cluster sur cette cartographie mettant en évidence une différence significative des réponses obtenues sur les électrodes postérieures.

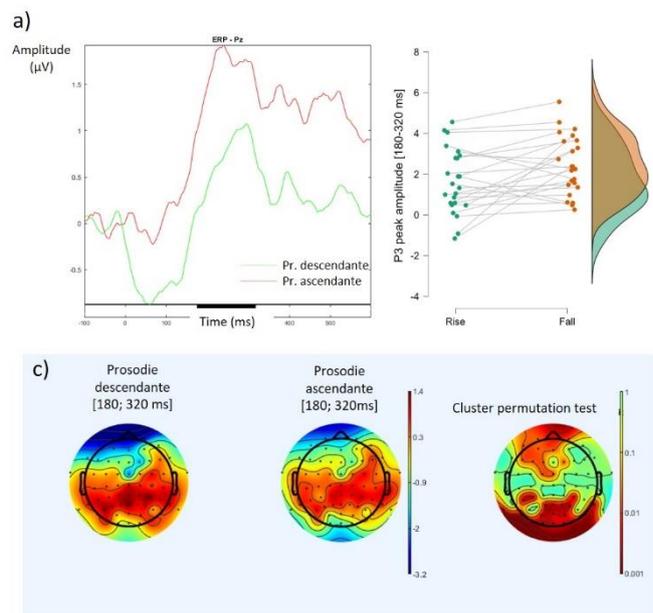


Figure 1. Comparaison des réponses électrophysiologiques obtenues suite à l'écoute du prénom prononcé avec une prosodie ascendante ou descendante

## Références bibliographiques

- André-Obadia, N., Zyss, J., Gavaret, M., Lefaucheur, J., Azabou, E., Boulogne, S., Guerit, J., McGonigal, A., Merle, P., Mutschler, V., Naccache, L., Sabourdy, C., Trébuchon, A., Tyvaert, L., Vercueil, L., Rohaut, B., & Yelnik, A. (2018). Recommendations for the use of electroencephalography and evoked potentials in comatose patients. *Neurophysiologie Clinique-clinical Neurophysiology*, 48(3), 143-169. <https://doi.org/10.1016/j.neucli.2018.05.038>
- Arias, P., Rachman, L., Liuni, M., & Aucouturier, J. (2021). Beyond Correlation: Acoustic Transformation Methods for the Experimental Study of Emotional Voice and Speech. *Emotion Review*, 13(1), 12-24. <https://doi.org/10.1177/1754073920934544>
- Pruvost-Robieux, E., André-Obadia, N., Marchi, A., Sharshar, T., Liuni, M., Gavaret, M., & Aucouturier, J. (2022). It's not what you say, it's how you say it : A retrospective study of the impact of prosody on own-name P300 in comatose patients. *Clinical Neurophysiology*, 135, 154-161. <https://doi.org/10.1016/j.clinph.2021.12.015>



# Flexibilité phonétique et Autisme

Eva GOESEELS<sup>1</sup>  
Virginie ROLAND<sup>1</sup>  
Myriam PICCALUGA<sup>1</sup>  
Clémence VERHAEGEN<sup>1</sup>  
Kathy HUET<sup>1</sup>  
Véronique DELVAUX<sup>1,2</sup>

<sup>1</sup> Service de Métrologie et Sciences du Langage  
Institut de Recherche en Sciences et Technologies du Langage,  
Université de Mons, Belgique

<sup>2</sup> Fond National de la Recherche Scientifique,  
Belgique

Notre thèse de doctorat concerne la flexibilité phonétique dans l'autisme. La flexibilité phonétique est définie comme l'aptitude du locuteur-auditeur à adapter son comportement de parole aux contraintes internes/externes pesant sur la situation de communication. Elle peut se traduire notamment par une tendance à la convergence phonétique en communication parlée, par une adaptation adéquate aux besoins de l'interlocuteur (personnes âgées, enfants, non-natifs), par une capacité à imiter la voix d'autrui/déguiser sa propre voix, par une aptitude à compenser une perturbation des conditions de communication, par une bonne maîtrise des sons/de la prosodie en langue étrangère, etc.

Chez les personnes neurotypiques, on considère que la flexibilité phonétique s'appuie sur divers fondements : fonctions exécutives (FE ; typiquement : flexibilité mentale), processus phonétiques (exploitation/actualisation des représentations sensori-motrices, boucles de rétroaction dans le contrôle moteur de la parole), compétences sociales et communicatives (Communication Accomodation Theory: Giles, 1987; Hordila-Vatamanescu & Pana, 2010). Chez les personnes avec autisme (dites TSA), certains de ces fondements pourraient être préservés/entraînés, du moins chez certains individus, d'autres sont typiquement déficitaires (FE : Demetriou et al., 2018 ; compétences communicatives : Eigsti et al., 2011; compétences sociales : Carter et al., 2005).

Kissine et al. (2021) ont étudié la flexibilité phonétique chez des adultes avec autisme via une tâche de compliance phonique, plus contrôlée qu'une tâche de parole spontanée (Kissine & Geelhand, 2019). Dans cette tâche empruntée à Delvaux et al. (2014), les 40 sujets (20 contrôles, 20 TSA) ont été amenés à produire les 10 voyelles du français (ligne de base de chaque sujet; LDB), et ensuite à reproduire 94 voyelles synthétiques non familières en langue maternelle, présentées auditivement via un casque. Les voyelles synthétiques ont été créées en faisant varier les trois premiers formants, pas-à-pas, avec ambition de remplir tout l'espace vocalique potentiel (Fig.1). Les auteurs ont mis en évidence une grande variabilité

dans les productions des adultes avec autisme, ainsi qu'une certaine « inflexibilité », mais se sont heurtés à des difficultés méthodologiques liées aux mesures de formants.

Dans cette communication, nous présentons nos questions de recherche et réflexions méthodologiques, ainsi qu'une analyse empirique fondée sur l'exploitation nouvelle des enregistrements recueillis par Kissine et al. (2021). En exploitant différentes métriques pour l'évaluation de la compliance (distance euclidienne moyenne par rapport à la cible entendue (DE) ; DE pondérée par la LDB des sujets ; variance des variances des productions par rapport à la cible), nous avons en effet pu mettre en évidence différents profils de flexibilité phonétique chez les adultes avec TSA (Fig.2).

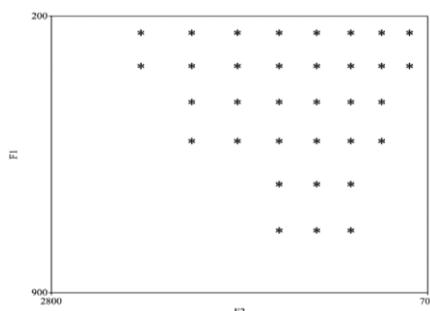


Figure 1. Représentation sur un plan F1-F2 de l'espace des stimuli présentés auditivement aux sujets.

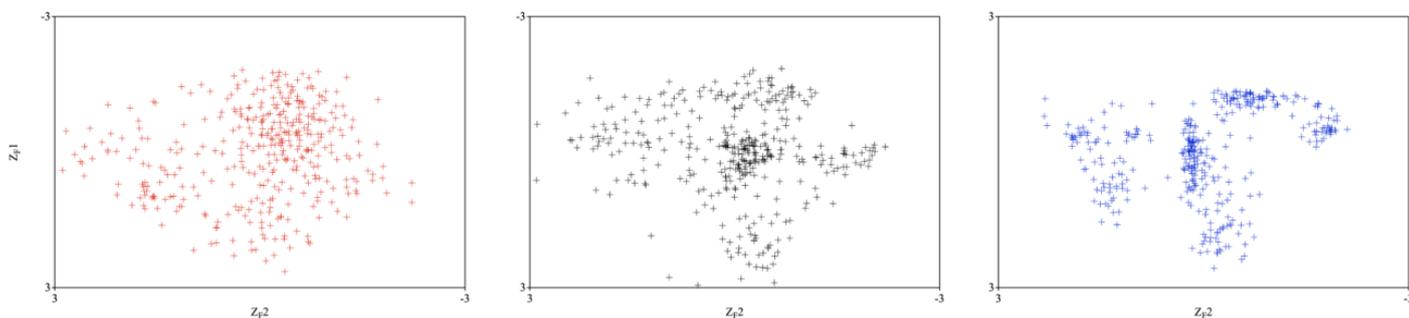


Figure 2. Productions vocaliques (F1, F2 : z-scores) de 3 locuteurs avec TSA (376 voyelles par participants) illustrant des profils différenciés en termes de répartition.

### Références bibliographiques

- Carter, A. S., Davis, N. O., Klin, A., & Volkmar, F. R. (2005). Social Development in Autism. In Volkmar, F.R., Paul, R., Klin, A., & Cohen, D. (Eds.), *Handbook of autism and pervasive developmental disorders: Diagnosis, development, neurobiology, and behavior*. 312–334.
- Delvaux, V., Huet, K., Piccaluga, M., & Harmegnies, B. (2014). Phonetic compliance: a proof-of-concept study. *Frontiers in psychology*, 5, 1375.
- Demetriou, E. A., Lampit, A., Quintana, D. S., Naismith, S. L., Song, Y. J. C., Pye, J. E., Hickie, I., & Guastella, A. J. (2018). Autism spectrum disorders: a meta-analysis of executive function. *Molecular psychiatry*, 23(5), 1198–1204.

- Eigsti, I.-M., de Marchena, A. B., Schuh, J. M., & Kelley, E. (2011). Language acquisition in autism spectrum disorders: A developmental review. *Research in Autism Spectrum Disorders*, 5(2), 681–691.
- Giles, H., Mulac, A., Bradac, J., & Johnson, P. (1987). *Speech accommodation theory: The first decade and beyond*. *Communication Yearbook*. London: McLaughlin, M.L., UK, Sage Publishers, 10, 13-48.
- Hordila - Vatamanescu, E.-M., & Pana, A.-D. (2010). The application of the Communication Accommodation Theory to virtual communities: A preliminary research on the online identities. *International Journal of Interdisciplinary Social Sciences*, 5(4), 279–290.
- Kissine, M., & Geelhand, P. (2019). Acoustic evidence for increased articulatory stability in the speech of adults with Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders*, 49, 2572-2580.
- Kissine, M., Geelhand, P., Philippart de Foy, M., Harmegnies, B., & Deliens, G. (2021), Phonetic Inflexibility in Autistic Adults. *Autism Research*, 14, 1186-1196.



# La mesure de l'intelligibilité en lecture par le biais de séquences délexicalisées chez des patients atteints de la maladie de Parkinson

Alexia BREVET<sup>1</sup>  
Manon JAYR<sup>1</sup>  
Alain GHIO<sup>1</sup>  
Muriel LALAIN<sup>1</sup>  
Gilles POUCHOULIN<sup>1</sup>  
Danielle DUEZ<sup>1</sup>  
Corinne FREDOUILLE<sup>2</sup>  
Virginie WOISARD<sup>3</sup>  
François VIALLET<sup>1</sup>

<sup>1</sup> Laboratoire Parole & Langage, UMR 7309 CNRS, Aix-Marseille Univ.

<sup>2</sup> Laboratoire Informatique d'Avignon, Univ. Avignon

<sup>3</sup> UT2J, Octogone-Lordat, Univ. & CHU Toulouse

La perte d'intelligibilité chez des patients atteints de troubles de la production de la parole est un élément important du bilan orthophonique. Il existe différents tests pour mesurer l'intelligibilité (Auzou & Rolland-Monnoury, 2006) mais nous estimons que ceux-ci sont biaisés et insuffisamment sensibles à cause de la restauration phonémique chez les auditeurs (Lalain et al, 2020). Nous proposons donc un test fondé sur des séquences délexicalisées de type Voyelle-Consonne-Voyelle (VCV) ainsi que sur des voyelles isolées (VOY) extraites automatiquement d'un corpus de lecture (Duez, Ghio & Viallet, 2020). Dans notre expérience, nous avons sélectionné 12 locuteurs contrôles et 30 patients en situation dopa ON/ dopa OFF. Pour chaque locuteur, nous avons extrait automatiquement 50 séquences VCV et 50 voyelles isolées. L'ensemble des 3600 séquences [50x(12+30x2)] ont été soumises à l'identification de la cible par 3 auditeurs différents. 18 auditeurs ont participé à l'expérience. L'expérience portait d'abord sur les 3600 consonnes puis sur les 3600 voyelles.

La mesure de l'intelligibilité est fondée sur le comptage du nombre de traits phonémiques mal perçus par l'auditeur (Perceived Phonological Deviation = PPD). Sur les consonnes, nos résultats montrent une différence significative entre le groupe contrôle (moyenne  $PPD_{\text{contrôle}}=0.31$ ) et les patients en situation de sevrage OFF dopa (moyenne  $PPD_{\text{parkOFF}}=0.69$ ) ou avec traitement ON dopa (moyenne  $PPD_{\text{parkON}}=0.72$ ). Nous n'observons aucun effet lié au traitement dopaminergique. En revanche, nous observons une augmentation du PPD en fonction de la sévérité de la dysarthrie évaluée cliniquement par le neurologue (*Figure 1*) :  $PPD_{\text{updrs0}}=0.54$  ;  $PPD_{\text{updrs1}}=0.63$  ;  $PPD_{\text{updrs2}}=0.86$  ;  $PPD_{\text{updrs3}}=1.15$ . Sur les voyelles, les résultats vont dans le même sens mais les différences entre les groupes sont

moins marquées ( $PPD_{CTRL}=0.86$  ;  $PPD_{updrs0}=0.95$  ;  $PPD_{updrs1}=1.05$  ;  $PPD_{updrs2}=1.26$  ;  $PPD_{updrs3}=1.32$ ).

Ces résultats valident notre méthode sur un corpus important, consolidant ainsi les résultats de Duez, Ghio & Viallet (2020). Nous confirmons aussi la pertinence des consonnes par rapport aux voyelles pour la mesure de l'intelligibilité de la parole. D'un point de vue clinique, nous n'observons pas d'effet lié au traitement pharmaceutique à la dopa. Dans le cadre du projet ANR Rugby<sup>14</sup>, nous avons appliqué cette méthode sur un corpus plus vaste de 316 locuteurs (111 CTRL +205 patients ON/OFF). Les données sont en cours de traitement. Bien qu'imparfait, le repérage automatique des séquences cibles reste efficace et cela nous conduit à imaginer, à terme, de proposer une version utilisable par les orthophonistes en cabinet pour tous les patients ayant des troubles de la production de la parole.

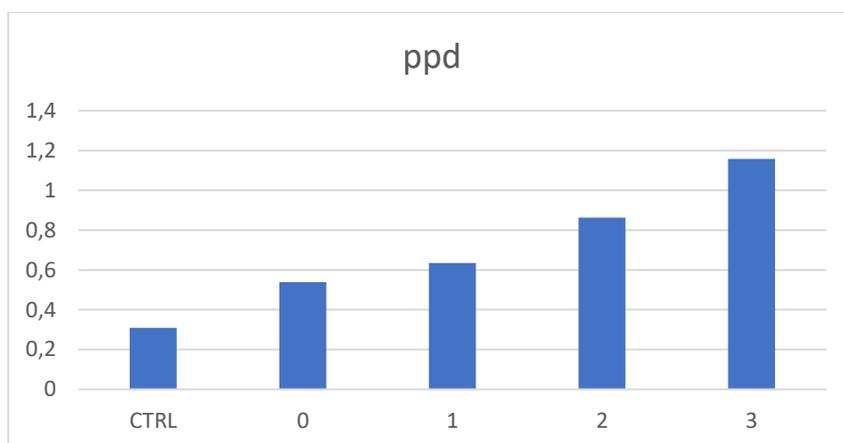


Figure 1. Déviation phonologique perçue (PPD) en fonction du grade de dysarthrie

### Références bibliographiques

- Auzou, P., & Rolland-Monnoury, V. (2006). *Batterie d'évaluation de la dysarthrie, 1st ed.* Isbergues: Ortho Edition.
- Duez, D., Ghio, A., & Viallet, F. (2020). Effect of linguistic context on the perception of consonants in Parkinsonian Read French speech. *Clinical Linguistics & Phonetics*, 35 (10), 926-944. <https://doi.org/10.1080/02699206.2020.1839969>
- Lalain, M., Ghio, A., Giusti, L., Robert, D., Fredouille, C., & Woisard, V. (2020). Design and Development of a Speech Intelligibility Test Based on Pseudowords in French : Why and How ? *Journal of Speech Language and Hearing Research*, 63(7), 2070-2083. [https://doi.org/10.1044/2020\\_jslhr-19-00088](https://doi.org/10.1044/2020_jslhr-19-00088)

<sup>14</sup> <https://anr.fr/Projet-ANR-18-CE45-0008> : Recherche d'unités linguistiques pertinentes pour améliorer la mesure de l'intelligibilité de la parole altérée par des troubles de production pathologique

# Comparaison de la structure formantique des voyelles en lecture et en parole spontanée dans la parole bègue et non bègue : une étude pilote

Laudy ARIDA<sup>1</sup>  
Ivana DIDIRKOVÁ<sup>2</sup>  
Fabrice HIRSCH<sup>1</sup>

<sup>1</sup> UMR 5267  
Praxiling & CNRS, Université Montpellier 3, Paul Valéry

<sup>2</sup> UR1569 Transcrit, Université Paris8  
Vincennes–Saint-Denis

Le bégaiement est défini comme « un trouble moteur de l'écoulement de la parole qui est alors produite avec plus d'effort musculaire, dans un contexte d'inversion du réflexe normal de décontraction... » (Monfrais-Pfauwadel, 2014). En tant que tel, il modifie alors l'exécution motrice des sons de la parole, notamment en situation de disflueuse (Didirková, Le Maguer & Hirsch, 2020). Parmi d'autres impacts sur la parole, des recherches se sont penchées sur la centralisation vocalique dans le bégaiement. Si certaines recherches ont montré une réduction de l'espace vocalique dans le bégaiement (Blomgren, 1998 ; Hirsch et al., 2007), ce n'est pas le cas d'autres études sur le sujet (Pendelieu-Verdurand et al., 2015).

Par ailleurs, il a été rapporté qu'il existe, dans la parole non pathologique, un lien entre réduction vocalique, débit et style de parole, avec une réduction de l'aire de triangle vocalique en parole spontanée (Meunier & Espesser, 2011). A notre connaissance, cette question reste peu documentée chez les sujets qui bégaiant.

C'est la raison pour laquelle nous avons choisi, dans cette étude, de mesurer les F1 et F2 des voyelles cardinales /a, i, u/ dans un groupe de personnes qui bégaiant (PQB) et un groupe de locuteurs de contrôle, afin de déterminer (1) si les valeurs de ces deux formants sont comparables dans les deux groupes et (2) si les PQB présentent une modification de ces deux formants dans la lecture et la parole spontanée.

Afin de répondre à ces questions, nous avons effectué des enregistrements audios de 18 locuteurs francophones dont 9 qui bégaiant et 9 qui ne bégaiant pas, appariés en sexe et en âge (5 hommes et 4 femmes âgés de 19 à 45 ans ( $\bar{x} = 26$  ans)), en parole semi-spontanée (réponses aux questions ouvertes, durée moyenne de l'enregistrement = 3,22 min) et en lecture (« Le pavillon d'Or »).

Les enregistrements ont été transcrits puis segmentés sur Praat (Boersma & Weenink, 2019). Un script ad-hoc y a été appliqué afin de relever les valeurs des formants sur le milieu (moyenne des valeurs relevées à 40, 50 et 60% de la durée totale) des 5312 voyelles

étudiées. Nous avons ensuite testé l'effet de la tâche et du groupe sur les valeurs formantiques.

Les résultats obtenus ne mettent en évidence ni un effet systématique de tâche ni un effet systématique de groupe. Les résultats seront présentés en détail lors de la conférence.

### Références bibliographiques

- Blomgren, M., Robb, M., & Chen, Y. (1998). A note on Vowel Centralization in Stuttering and Nonstuttering Individuals. *Journal of Speech, Language, and Hearing Research*, 41, 1042-1051.
- Boersma, P., & Weenink, D. (2019). Praat: Doing phonetics by computer (version 6.1.01). <http://www.praat.org>.
- Didirková, I., Le Maguer, S., & Hirsch, F. (2020). An articulatory study of differences and similarities between stuttered disfluencies and non-pathological disfluencies. *Clinical Linguistics & Phonetics*, 1–21. doi:10.1080/02699206.2020.1752803.
- Hirsch, F., Fauvet, F., Ferbach-Hecker, V., Béchet, M., Bouarourou, F., & Sturm, J. (2007). Formant structures of vowels produced by stutterers at normal and fast speech rates. ICPHS 2007 proceedings, 1345-1348.
- Meunier, C., & Espesser, R. (2011). Vowel reduction in conversational speech in French : The role of lexical factors. *Journal of Phonetics*, 39(3), 271-278. doi: 10.1016/j.wocn.2010.11.008.
- Monfrais-Pfauwadel, M.-C. (2014). *Bégaiement, bégaiements : un manuel clinique et Thérapeutique*, Paris : De Boeck-Solal.
- Pendelieu-Verdurand, M., Tissot-Dupont, A., Rossato, S., & Zmarich, C. (2015). Coarticulation and vowel centralization in French and Italian stuttering adults and children. *Second International Conference on Stuttering*, 204-206.

# De la valeur aberrante à l'altération *rare* de la parole : application à la parole pathologique.

Richard JU DENNE<sup>1</sup>  
Timothy POMMEE<sup>2,3</sup>  
Régine ANDRE-OBRECHT<sup>1</sup>  
Julie MAUCLAIR<sup>1</sup>

<sup>1</sup> Institut de Recherche en Informatique de Toulouse,  
UMR 5505 CNRS,  
Université Toulouse 3 – Paul Sabatier

<sup>2</sup> Bulbar Function Lab, Sunnybrook Research Institute

<sup>3</sup> Speech Production Lab, University of Toronto

« Mesurer » si une parole est *normale*, *atypique* voire *pathologique* relève du défi. Après apprentissage statistique de modèles, une mesure de similarité (Quintas et al., 2022) évalue le *comportement moyen* de la parole du locuteur, ignorant de fait les quelques altérations « rares » de sa production, assimilées à des « données statistiquement aberrantes » alors qu'elles pourraient être révélatrices de dysfonctionnement.

Pour illustrer cette réflexion, une expérience est menée dans le cadre du projet ANR Voice4PD-MSA<sup>15</sup> où est étudiée la parole de patients atteints de la maladie de Parkinson (PD) ou d'Atrophie Multi-Systématisée (AMS). La parole de ceux diagnostiqués à un stade précoce de la maladie reste « normale », en témoignent les mesures de sévérité élevées (8 à 10) (Woisard et al., 2013) ; la différenciation entre locuteurs sains et patients est difficile. D'où la recherche d'altérations rares lors de prononciations répétées de /pataka/, sachant que les plosives peuvent alors être mal prononcées (Das., 2021) et que l'interprétation clinique est possible compte tenu de leur caractérisation acoustico-articulatoire (Stevens et al., 1978).

Pour ce faire, sur la zone de relâchement de l'occlusion des plosives /t/ et /k/, détectées automatiquement, sont extraits le centre de gravité spectral (cog), la variance spectrale (std) et l'asymétrie spectrale (skew) du spectre d'énergie calculé sur 5ms (Pommée, 2021). Le nuage de points associés à chaque patient est visualisé, ainsi que le nuage « sain » formé de l'ensemble des extractions des locuteurs sains, par projection dans l'espace (cog, std, skew). Sont étudiés 3 patients « PD » et 3 « MSA ». Les *Figures 1* et *2* illustrent cette représentation pour un patient « PD1 » et un patient « MSA1 ». L'observation des « outliers » au sens statistique montrent que :

- Leur nombre est plus conséquent pour le seul locuteur « MSA1 », dans le plan (cog, std) du /t/, et pour le seul locuteur « PD1 », dans le plan (cog, skew) du /k/.
- Le plan (cog, skew) du /t/ en révèle pour les deux locuteurs, mais différemment.

---

<sup>15</sup> <https://anr.fr/Projet-ANR-16-CE19-0010>

Au vu de cette illustration, il convient d'étendre le nuage associé aux sujets sains et d'observer une cohorte plus importante de patients. Après confirmation, il faudra rechercher une formalisation de l'approche : comment mettre en évidence les « outliers » ? Comment les interpréter pour non pas aboutir à un diagnostic clinique, mais une simple alerte ?

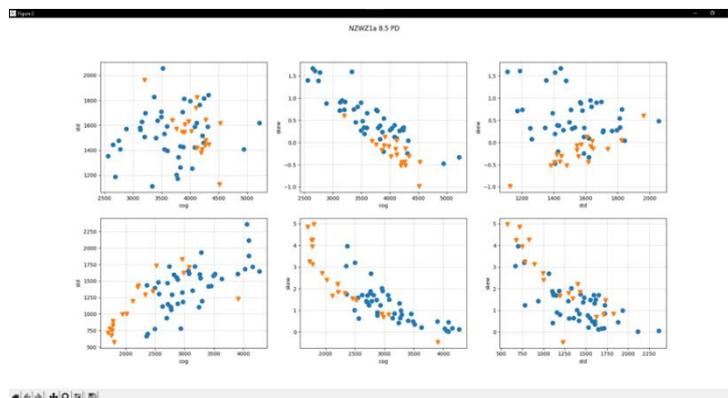


Figure 1. Visualisation des moments spectraux dans les trois plans pour le phonème /t/ (ligne supérieure) et le phonème /k/ (ligne inférieure). Le nuage « sain » est représenté en bleu tandis que celui du patient « PD1 » est en orange (sévérité=8.5).

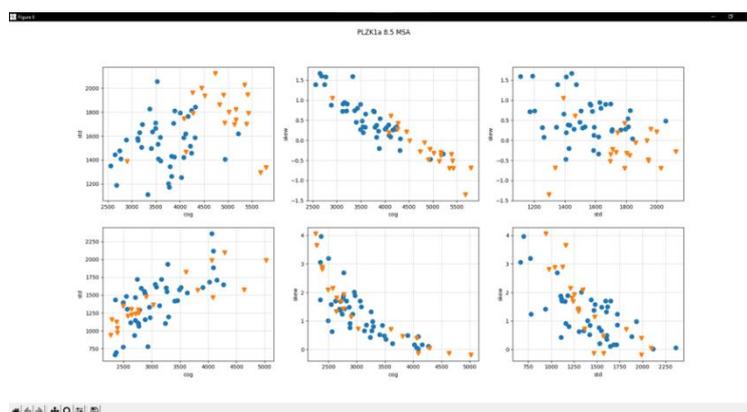


Figure 2. Visualisation des moments spectraux dans les trois plans pour le phonème /t/ (ligne supérieure) et le phonème /k/ (ligne inférieure). Le nuage « sain » est représenté en bleu tandis que celui du patient « MSA1 » est en orange (sévérité=8.5).

## Références bibliographiques

- Das, B., (2021). *Marqueurs vocaux objectifs automatisés pour le diagnostic différentiel entre la maladie de Parkinson et les troubles parkinsoniens atypiques*, Thèse de l'université de Bordeaux, septembre 2021.
- Pommée, T., (2021). *Les mesures d'intelligibilité : état de l'art, considérations pratiques pour l'applicabilité clinique et explorations acoustiques*, Thèse de l'université Toulouse 3 – Paul Sabatier, novembre 2021.

- Quintas, S., Mauclair, J., Woisard, V., & Piquier, J. (2022), Automatic Assessment of Speech Intelligibility using Consonant Similarity for Head and Neck Cancer. *Proc. of Interspeech 2022*, South Korea.
- Stevens, K. N., & Blumstein, S. E. (1978). Invariant cues for place of articulation in stop consonants. *Journal of the Acoustical Society of America*, 64(5), 1358-1368. <https://doi.org/10.1121/1.382102>
- Woisard, V., Espesser, R., Ghio, A., & Duez, D. (2013), De l'intelligibilité à la compréhension de la parole, quelles mesures en pratique clinique ? *Revue de Laryngologie Otologie Rhinologie*, 134(1), 27-33, 2013.



# Analyse des performances de systèmes de reconnaissance automatique de la parole spontanée après cancer oral ou oropharyngé

Mathieu BALAGUER<sup>1</sup>

Julien PINQUIER<sup>1</sup>

Virginie WOISARD<sup>2,3</sup>

Jérôme FARINAS<sup>1</sup>

<sup>1</sup> IRIT, Université de Toulouse, CNRS, Toulouse INP, UT3, Toulouse, France

<sup>2</sup> Hôpital Larrey, Hôpitaux de Toulouse, France

<sup>3</sup> Laboratoire de Neuro-Psycho-Linguistique LNPL, Université Toulouse II, France

Contexte : L'analyse de la parole spontanée après cancer est essentielle en clinique, car il s'agit de la situation de production la plus « écologique » (Knuijt et al., 2017, Prins & Bastiaanse, 2004). Elle nécessite de pouvoir s'appuyer sur une transcription fiable. Actuellement, les systèmes de reconnaissance automatique de parole facilitant la transcription sont entraînés sur la parole typique, et prennent mal en compte les spécificités de la parole après cancer (Balaguer et al., 2021). Or, une reconnaissance lexicale plus juste permettrait des analyses plus précises du discours des patients, et l'accès à de nouvelles informations sur la dynamique psychosociale des patients.

Objectifs : Étudier les performances de systèmes de reconnaissance automatique de parole (RAP) au niveau du mot appliqués à la parole spontanée après cancer.

Méthodes : La parole spontanée de 12 patients traités pour un cancer oral ou oropharyngé a été enregistrée au cours d'un entretien. Une minute a été transcrite au niveau orthographique par un orthophoniste expert et deux sujets naïfs, sur l'hypothèse qu'un système entraîné sur de la parole typique aurait un comportement plus proche d'un auditeur naïf par réduction des effets d'expertise (Fex, 1992). Une transcription automatique au niveau du mot a été réalisée par plusieurs systèmes de RAP (*Tableau 1*). La boîte à outils HTK (Young et al., 2015) a permis d'obtenir des scores de comparaison avec les transcriptions des experts puis avec les naïfs. Des tests de comparaison 2 à 2 (test de Wilcoxon) ont été menés pour comparer les performances des systèmes entre eux, puis les performances de chaque système selon l'expertise du juge humain (expert *vs* naïfs).

Tableau 1. Systèmes de RAP étudiés

Système	Architecture	Corpus d'entraînement (heures)
CRDNN	End-to-end Convolutional Recurrent Deep Neural Network (Heba et al., 2021)	LibriSpeech (960 h)
TDDNf	Time-Delay Neural Network factorisé – Hidden Markov Model (Gelin et al., 2021)	CommonVoice (147 h)
Whisper (Open AI)	Transformer Seq2seq (Radford et al., 2022)	Multilingue (117.113 h)

Résultats : Whisper présente une exactitude significativement meilleure ( $p=0,004$ ) que le TDNNf bien que restant faible (taux d'erreurs mots = 76,47 % par rapport à l'expert !), et un pourcentage de mots corrects, un nombre de mots corrects et de délétions significativement meilleurs que les deux autres systèmes ( $p<0,05$ ), tant en comparaison avec l'expert qu'avec les naïfs. En revanche, le nombre d'insertions et de substitutions est significativement plus élevé que pour le CRDNN (Tableau 2).

Tableau 2. Comparaison des systèmes de RAP selon le profil expert/naïf de l'examineur (en gras, la meilleure valeur par métrique)

	Expert			Naïfs		
	CRDNN	TDNNf	Whisper	CRDNN	TDNNf	Whisper
Exactitude $Ex = \frac{100 \times (H-I)}{H+S+D}$	21,17 %	13,79 %	<b>23,53 %</b>	21,07 %	13,40 %	<b>27,43 %</b>
Pourcentage de mots corrects $P_{Corr} = \frac{100 \times H}{H+S+D}$	21,88 %	16,33 %	<b>37,87 %</b>	22,11 %	16,33 %	<b>36,44 %</b>
Nombre de mots correctement reconnus (H)	20	14,33	<b>32,83</b>	18,83	13,54	<b>29,42</b>
Nombre de délétions (D)	39,83	26,67	<b>15,92</b>	32,58	22,04	<b>10,25</b>
Nombre de substitutions (S)	<b>18,33</b>	37,17	29,42	<b>18,42</b>	37,50	31
Nombre d'insertions (I)	<b>0,5</b>	1,75	8	<b>0,75</b>	2,21	9,83

Les systèmes présentent des performances globalement similaires, que leur référence de comparaison soit un expert ou les naïfs. Seul Whisper semble avoir une transcription plus proche de celle des naïfs avec une exactitude significativement plus élevée comparativement aux experts (Tableau 3).

Tableau 3. Comparaison des performances de chaque système de RAP selon le profil expert/naïf de l'examinateur (p-value, test de Wilcoxon : en gras, les p-values significatives au seuil de 5 %, les valeurs de chaque système sont à retrouver dans le Tableau 2)

Système	CRDNN	TDNNf	Whisper
Exactitude	0,84	0,16	<b>0,049</b>
Pourcentage de mots corrects	0,5	0,75	0,18
Mots corrects	<b>0,02</b>	<b>0,02</b>	<b>0,003</b>
Délétions	<b>0,02</b>	<b>0,02</b>	<b>0,004</b>
Substitutions	0,28	0,24	0,14
Insertions	0,16	0,66	0,11

Conclusion : Cette étude exploratoire montre que l'architecture des systèmes de RAP et la taille de leur corpus d'entraînement (sur parole typique) influencent les performances de reconnaissance de la parole après cancer. Toutefois, l'exactitude reste faible. Cela montre la nécessité d'adapter les modèles acoustiques à la parole cancérologique pour permettre des analyses plus précises du discours informant sur la dynamique psychosociale des patients et l'ajustement des stratégies thérapeutiques. Ces résultats préliminaires devront être étudiés sur un échantillon plus large, notamment pour analyser l'effet de la sévérité du trouble de parole sur les performances de reconnaissance.

### Références bibliographiques

- Balaguer, M. (2021). *Mesure de l'altération de la communication par analyses automatiques de la parole spontanée après traitement d'un cancer oral ou oropharyngé*. Thèse de doctorat, Université Paul Sabatier Toulouse III. <https://theses.hal.science/tel-03557511>
- Fex, S. (1992). Perceptual evaluation. *Journal of Voice*, 6(2), 155-158
- Gelin, L., Daniel, M., Pinquier, J., & Pellegrini, T. (2021). End-to-end acoustic modelling for phone recognition of young readers. *Speech Communication*, 134, 71–84. <https://doi.org/10.1016/j.specom.2021.08.003>
- Heba, A. (2021). *Reconnaissance automatique de la parole à large vocabulaire : des approches hybrides aux approches End-to-End*. Thèse de doctorat, Université Paul Sabatier Toulouse III. <https://tel.archives-ouvertes.fr/tel-03616588>
- Knuijt, S., Kalf, J. G., van Engelen, B. G. M., de Swart, B. J. M., & Geurts, A. C. H. (2017). The Radboud Dysarthria Assessment: Development and Clinimetric Evaluation. *Folia Phoniatrica et Logopaedica*, 69(4), 143–153. <https://doi.org/10.1159/000484556>
- Prins, R., & Bastiaanse, R. (2004). Analysing the spontaneous speech of aphasic speakers. *Aphasiology*, 18(12), 1075–1091. <https://doi.org/10.1080/02687030444000534>

- Radford, A., Wook, J., Tao, K., Greg, X., Christine, B., & Ilya, M. (2021). Robust Speech Recognition via Large-Scale Weak Supervision. *Openai.Com*.  
<https://doi.org/10.48550/arXiv.2212.04356>
- Young, S., Evermann, G., Gales, M., Hain, T., Kershaw, D., Liu, X. (Andrew), Moore, G., Odell, J., Ollason, D., Pover, D., Ragni, A., Valtchev, V., Woodland, P., & Zhang, C. (2015). *The HTK Book (for HTK Version 3.5, documentation alpha version) (Issue December)*. Cambridge University Engineering Department.

# Diagnostic différentiel chez les enfants avec un trouble des sons de parole : intérêt de l'épreuve de diadococinésies

Noémie WINLING<sup>1</sup>  
Geneviève MELONI<sup>2</sup>  
Andrea A. N. MacLEOD<sup>3</sup>

<sup>1</sup> Centre de Formation Universitaire en Orthophonie,  
Université de Besançon

<sup>2</sup> Faculty of Medicine and Health Sciences, School of Communication Sciences and  
Disorders,  
McGill University

<sup>3</sup> Faculty of Rehabilitation Medicine, Communication Sciences & Disorders,  
University of Alberta

Les troubles des sons de parole (TSP) correspondent à un délai de développement des sons de la langue conduisant à une perte d'intelligibilité (Baker & McLeod, 2011). Ces troubles sont un enjeu pour les orthophonistes car leurs conséquences sont importantes sur la socialisation et le devenir scolaire des enfants (Felsenfeld, Broen et McGue, 1994). Les enfants avec un TSP peuvent présenter une combinaison de difficultés au niveau de la perception de la parole, de la production et/ou des représentations phonologiques (IEPMCS, 2012). Les profils de TSP sont très hétérogènes : certains enfants peinent à structurer leurs représentations phonologiques et d'autres présentent des déficits au niveau de la programmation et planification motrice de la parole. Identifier des sous-types de TSP permet donc de choisir une rééducation adaptée.

Pour aider au diagnostic des enfants avec un TSP, des chercheurs et des orthophonistes se sont rassemblés pour créer l'outil d'évaluation EULALIES qui comprend une tâche de diadococinésies. Les diadococinésies se définissent comme la capacité à réaliser rapidement des mouvements alternatifs, comme /pa/, /ta/ et /ka/. Elles permettent de mettre en lumière des difficultés au niveau de la programmation motrice (Kent et al., 2022). Bien que cette tâche semble être utile et facile d'administration, il y a présentement peu de recherche sur les enfants francophones (Kent et al., 2022). La présente étude adresse ce manque.

Nous analysons l'effet de l'âge, du diagnostic, du statut linguistique et du genre sur les performances aux séries diadococinésiques chez 47 enfants franco-québécois âgés de 4 à 8 ans. Nous exploitons 3 types de mesures : un score de vitesse, c'est-à-dire le nombre de syllabes produites en 10 secondes, un score de constance, qui correspond au nombre de versions différentes dans les cinq premières répétitions et un score de précision, qui reflète le nombre d'erreurs.

Les résultats de cette étude montrent un effet de l'âge sur la vitesse : les enfants plus âgés produisent davantage de répétitions que les enfants les plus jeunes. Nous retrouvons également un effet du diagnostic sur la vitesse : les enfants présentant une dyspraxie verbale sont significativement plus lents que les autres enfants. Enfin nous notons un effet du diagnostic sur le score de constance, les enfants avec un Trouble des Sons de la Parole sont plus inconstants que les autres.

Les diadococinésies apparaissent donc comme un outil simple et puissant d'évaluation des compétences articulatoires. Elles peuvent aider au diagnostic des enfants avec un TSP.

### Références bibliographiques

- Baker, E., & McLeod, S. (2011). Evidence-Based Practice for children with speech sound disorders: Part 1 Narrative Review, *Language, Speech, and Hearing Services in Schools*, 42, 102-139. DOI:10.1044/0161-1461(2010/09-0075)
- Felsenfeld, S., Broen, P. A., & McGue, M. (1994). A 28-year follow-up of adults with a history of moderate phonological disorder: Educational and occupational results. *Journal of Speech and Hearing Research*, 37(6),1341-1353. DOI:10.1044/jshr.3706.1341
- International Expert Panel on Multilingual Children's Speech (2012). *Multilingual children with speech sound disorders: Position paper*. Bathurst, NSW, Australia: Research Institute for Professional Practice, Learning and Education (RIPPLE), Charles Sturt University. Repéré à <http://www.csu.edu.au/research/multilingual-speech/position-paper>
- Kent, R. D., Kim, Y., & Chen, L. M. (2022). Oral and Laryngeal Diadochokinesis Across the Life Span: A Scoping Review of Methods, Reference Data, and Clinical Applications. *Journal of Speech, Language, and Hearing Research*, 65(2), 574-623.

# Coarticulation anticipatoire V-à-V dans différentes dysarthries

Daria D'ALESSANDRO<sup>1</sup>  
Angéline BOURBON<sup>1</sup>  
Michaela PERNON<sup>1,2</sup>  
Marina LAGANARO<sup>3</sup>  
Frédéric ASSAL<sup>4</sup>  
Cécile FOUGERON<sup>1</sup>

<sup>1</sup>Laboratoire de Phonétique et Phonologie, UMR 7018 CNRS/Université Sorbonne  
Nouvelle-Paris 3

<sup>2</sup> CRMR Wilson & Unité Parkinson, Service de Neurologie, Hôpital Fondation A.  
de Rothschild, Paris

<sup>3</sup> Laboratoire de Psycholinguistique,  
Équipe Neuro-psycho-linguistique, FPSE,  
Université de Genève

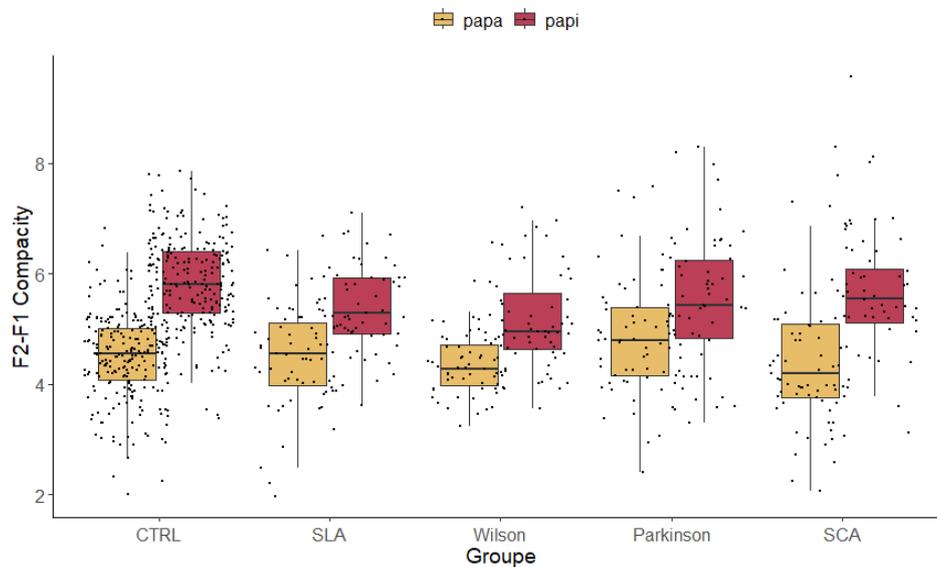
<sup>4</sup> Hôpitaux Universitaires de Genève, Département des  
Neurosciences cliniques

Une altération de la coarticulation est décrite comme une des possibles caractéristiques des troubles moteurs acquis de la parole (e.g. Ziegler & Von Cramon, 1986 ; Hertrich & Ackermann, 1999), et pourrait être attribuée à différents facteurs, tels qu'un ralentissement de la parole, une amplitude réduite des mouvements, ou une difficulté dans la planification et coordination de mouvements successifs (e.g. Duffy, 2012). Cependant, si une atteinte de la coarticulation dans l'apraxie de la parole est plus systématiquement montrée, les études portant sur la coarticulation dans la dysarthrie, se focalisant surtout sur la coarticulation voyelle-à-consonne, rapportent des résultats inconsistants (e.g. Tjaden, 2000 ; Martel-Savageau & Tjaden, 2017).

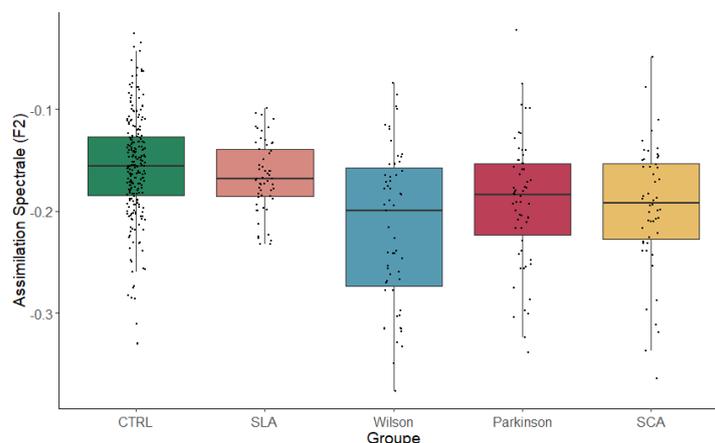
Dans cette étude, la coarticulation anticipatoire voyelle-à- voyelle est étudiée chez des locuteurs atteints par différents types de dysarthries associées à la Sclérose Latérale Amyotrophique (SLA), la maladie de Wilson, la maladie de Parkinson et l'Ataxie Spino-Cérébelleuse (SCA).

La production de 40 locuteurs dysarthriques est comparée avec la production de 40 contrôles appariés en sexe et âge (*Tableau 1*). La coarticulation est considérée comme l'influence de V2 /i/ sur V1 /a/ dans le mot *papi*, vs. V2 /a/ dans *papa*. Les deux mots sont insérés dans un court texte dont la lecture fait partie du protocole MonPaGe, conçu pour l'évaluation de la parole pathologique (Pernon et al., 2020). Premièrement, la coarticulation est analysée comme la différence entre V1 de *papi* vs *papa*, calculée sur une mesure de compacité F2-F1. Deuxièmement, l'assimilation acoustique de V1 /a/ à V2 /i/ dans *papi* est calculée sur les deux dimensions F1 et F2 séparément, en utilisant la formule suivante  $((F[V1a] - F[V2i]) / F[V2i])$ .

Si tous les groupes présentent une différence significative entre la compacité F2-F1 de *papi* vs *papa*, et donc de la coarticulation V-à-V, cette différence est réduite chez les locuteurs dysarthriques (*Figure 1, Tableau 2*). Les résultats sur l'assimilation acoustique de V1/a/ à V2/i/ montrent une réduction significative de la coarticulation pour les groupes Wilson, Parkinson et SCA sur la dimension F2 (*Figure 2, Tableau 3*) par rapport aux contrôles. Cependant, aucune différence n'est trouvée sur F1. Des analyses ultérieures relèvent que ce résultat peut être attribué à la production de /a/ plus fermées chez les dysarthriques dans les deux contextes (V2 /a/ et /i/), et donc des /a/ plus proches de /i/ sur cette dimension. Cela peut être interprété en lien avec une amplitude réduite des mouvements de la mâchoire, ou bien avec une plus grande coarticulation de la voyelle avec la bilabiale /p/.



*Figure 1.* Compacité F2-F1 pour V1 /a/ de *papa* (en jaune) et *papi* (en rouge), pour tous les groupes : contrôles (CTRL), Sclérose Latérale Amyotrophique (SLA), maladie de Wilson (Wilson), maladie de Parkinson (Parkinson), Ataxie Spino-Cérébelleuse (SCA)



*Figure 2.* Assimilation spectrale de /a/ à /i/ dans *papi* sur la dimension F2 pour tous les groupes : contrôles (CTRL), Sclérose Latérale Amyotrophique (SLA), maladie de Wilson (Wilson), maladie de Parkinson (Parkinson), Ataxie Spino-Cérébelleuse (SCA).

Tableau 1. Résumé des caractéristiques des locuteurs par groupe : nombre de locuteurs pour chaque sexe, âge (moyenne et déviation standard) et type de dysarthrie.

Groupe	Nombre	Age	Type de Dysarthrie
SLA	7F, 3H	M=74, SD=4.4	Mixte (Spastique/Flaccide)
Wilson	1F, 9H	M=35.5, SD=7.7	Mixte (Hypo/Hyperkinétique)
Parkinson	3F, 7H	M=72.5, SD=11	Hypokinétique
SCA	5F, 5H	M=55.9, SD=8.3	Ataxique
CTRL	16F, 24H	M=59, SD=17	/

Tableau 2. Effet de V2 (/i/ ou /a/ sur la compacité F2-F1 de V1 /a/ dans *papa* vs *papi*. Chi carré, degrés de liberté, valeur de p pour les effets fixes et les interactions.

Effet de V2 et du groupe sur la compacité F2-F1 de V1 /a/ Résumé des effets fixes et des interactions ( $R^2_m=0,26$ ; $R^2_c=0,77$ )	
V2 (/a/ vs /i/)	$\chi^2(1)=65,5$ ; $p<0,0001$
Groupe	<i>ns</i>
V2*Groupe	$\chi^2(4)=15,6$ ; $p=0,004$

Tableau 3. Effet du groupe sur l'assimilation acoustique entre F2 de V1/a/ et F2 de V2/i/. Pour chaque groupe, coefficient bêta, erreur standard ; valeur de t et valeur de p.

Effet du groupe sur l'assimilation acoustique (F2) de V1 /a/ à V2 /i/ Résumé des effets fixes ( $R^2_m=0,13$ ; $R^2_c=0,68$ )	
SLA	$\beta=-0,01$ ; $SE=0.008$ ; $t=-19,47$ ; $p=ns$
Wilson	$\beta=-0,06$ ; $SE=0.017$ ; $t=-3,66$ ; $p=0,0005$
Parkinson	$\beta=-0,03$ ; $SE=0.017$ ; $t=-2,1$ ; $p=0,04$
SCA	$\beta=-0,04$ ; $SE=0.017$ ; $t=-2,3$ ; $p=0,02$

## Références bibliographiques

- Duffy, J. R. (2012). *Motor speech disorders: Substrates, differential diagnosis, and management*. Mosby, 3rd edition. 2012.
- Hertrich, I., & Ackermann, H. (1999). Temporal and spectral aspects of coarticulation in ataxic dysarthria: An acoustic analysis. *Journal of Speech, Language, and Hearing Research*, 42(2), 367-381. 1999.
- Martel-Sauvageau, V., & Tjaden, K. (2017). Vocalic transitions as markers of speech acoustic changes with STN-DBS in Parkinson's Disease. *Journal of Communication Disorders*, 70, 1-11. <https://doi.org/10.1016/j.jcomdis.2017.10.001>
- Pernon, M., Leveque, N., Delvaux, V., Assal, F., Borel, S., Fougeron, C., Trouville, R., & Laganaro, M. (2020). MonPaGe, un outil de screening francophone informatise d'évaluation perceptive et acoustique des troubles moteurs de la parole (dysarthries, apraxie de la parole). *Rééducation orthophonique*, 281, 171-97.
- Tjaden, K. (2000). An Acoustic Study of Coarticulation in Dysarthric Speakers With Parkinson Disease. *Journal of Speech Language and Hearing Research*, 43(6), 1466-1480. <https://doi.org/10.1044/jslhr.4306.1466>
- Ziegler, W., & Von Cramon, D. (1986). Disturbed coarticulation in apraxia of speech: Acoustic evidence. *Brain and Language*, 29(1), 34-47. 1986.

# Similitudes entre les mécanismes de la déglutition et la production de la parole

Marta P. CIRCIU<sup>1,2</sup>  
Didier DEMOLIN<sup>1</sup>

<sup>1</sup>Laboratoire de Phonétique et Phonologie,  
Université Sorbonne Nouvelle -Paris 3/ CNRS-UMR 7018

<sup>2</sup>Hôpital Foch, Service de Laryngologie Phoniatrie, Suresnes  
Université de Versailles Saint-Quentin en Yvelines

Cette communication propose de montrer que certaines classes de sons de la parole impliquent le réemploi de mécanismes de la déglutition. Les conclusions de cette contribution résultent des données cliniques en radiocinéma, d'IRM en temps réel, de mesures aérodynamiques et d'acoustique.

La phase orale, préparatoire et propulsive est le temps volontaire de la déglutition. La phase pharyngée qui suit est un temps rapide et complexe de la déglutition pour propulser le bolus (liquide ou solide) à travers le pharynx, vers l'œsophage tout en protégeant les voies aériennes supérieures (Cicero & Murdoch, 2006).

Ces deux phases montrent deux intervalles de temps où la pression de l'air dans la cavité supra-glottique est changée en modifiant son volume. Pendant la phase préparatoire, le volume entre une fermeture linguale antérieure et postérieure est amplifié, ce qui diminue la pression de l'air dans cette cavité. Pendant la phase pharyngée, le volume situé entre le port vélo-pharyngé et l'additus laryngé fermés est contracté, responsable d'une augmentation de la pression de l'air dans cette cavité. Ces deux manœuvres sont similaires à celles qui sont impliquées dans la production des clics et des consonnes éjectives (Hixon et al., 2020).

L'expansion de la cavité orale entre les occlusions antérieure et postérieure ont été observées, en radiocinéma et en IRM temps réel pendant la déglutition et la production de la parole. Dans la phase préparatoire orale, et contrairement aux étapes propulsives orale et pharyngée, les changements de pression de l'air dans ces cavités sont la cible des variations de volume, lorsque ces manœuvres sont réutilisées pour parler, en particulier pour produire les sources vocales caractéristiques des clics et des éjectives (Traill, 1987).

Une condition nécessaire à la production de toute source sonore vocale est une différence de pression entre une cavité et une autre dans le conduit vocal. Le son est produit par le flux d'air passant à travers la valve reliant les deux cavités. La classe de sons qui est pertinente ici est caractéristique des obstruantes. Les variations de pression d'air ont été observées lors de mesures de pression intra-orale pendant la déglutition et les consonnes labio-dorsales.

Ces observations basées sur des données d'IRM et temps réel, de radiocinéma et des mesures de pression intra-orale, permettent d'étayer l'hypothèse de la similitude des mécanismes de la déglutition et la production de certaines classes de sons de la parole.

## Références bibliographiques

- Cicero, J. A. Y., & Murdoch, B.E. (2006). *Dysphagia: Foundations, Theory and Practice*. New-York: Wiley.
- Hixon, T.J., Weismer, G., & Hoit, J.D. (2020). *Preclinical Speech Science*. 3<sup>rd</sup> edition. San Diego: Plural Publishing.
- Traill, A. (1997). Linguistic phonetic features for clicks: Articulatory, acoustic and perceptual evidence. In R.K. Herbert (ed.) *African Linguistics at the Crossroads: Papers from Kwalusemi*. Köln: Rudiger Köppe, 97-117

# Classification d'Alzheimer à partir de paramètres acoustiques et prosodiques avec de l'apprentissage automatique

Eric JORDAN<sup>1</sup>  
Jalal AL-TAMIMI<sup>1</sup>

<sup>1</sup> Laboratoire de Linguistique Formelle,  
Université de Paris Cité

La maladie d'Alzheimer est difficile à diagnostiquer car elle nécessite soit des procédures invasives soit des examens du patient par des spécialistes. Des études montrent qu'intervenir tôt dans le développement de cette maladie serait plus efficace pour améliorer la qualité de vie du patient (De Roeck et al., 2019). Afin d'optimiser et accélérer le diagnostic, il serait intéressant de développer des méthodes automatisées (De La Fuentes Garcia, Richie & Luz, 2020). C'est pourquoi, cette étude évalue la capacité des systèmes d'apprentissage automatique sur plusieurs ensembles de mesures acoustiques et prosodiques.

Les données utilisées sont des enregistrements de parole spontanée (description de l'image « Cookie Theft » ; une tâche souvent utilisée dans le diagnostic de la maladie d'Alzheimer) appartenant à un ensemble de données créé pour l'ADReSS Challenge (Luz et al., 2020), un défi qui a eu lieu à Interspeech 2020, avec l'objectif de développer des solutions automatisées pour le diagnostic de la maladie d'Alzheimer. Au total, il y a 156 participants (86 F, 70 H) âgés entre 50 et 80 ans. Un avantage de ces données est l'équilibre des participants non atteints (non-AD) et atteints (AD). Aucune information précise concernant le stade d'avancement de la pathologie n'est fourni, cependant les résultats du test MMSE fournis indiquent des stades d'avancement variés. Ces données sont accessibles via la DementiaBank (Becker et al., 1994).

Des méthodes de traitement automatisées ont été appliquées à ces enregistrements afin d'extraire des mesures. Une segmentation automatique a été réalisée avec le logiciel Prosogram (Mertens, 2004). Trois ensembles de mesures acoustiques ont été automatiquement obtenus avec les logiciels Prosogram et OpenSmile (Eyben, Wöllmer & Schuller, 2010), via Python. Ces ensembles furent : GeMAPS (Eyben et al., 2015) (jusqu'à 62 mesures spectrales, temporelles, fréquence, énergie, qualité vocale), ComParE (Schuller et al., 2016) (jusqu'à 6373 mesures, spectrales, cepstrales, temporelles, fréquence, énergie, qualité vocale) et Prosogram (41 mesures, temporelles, et fréquence).

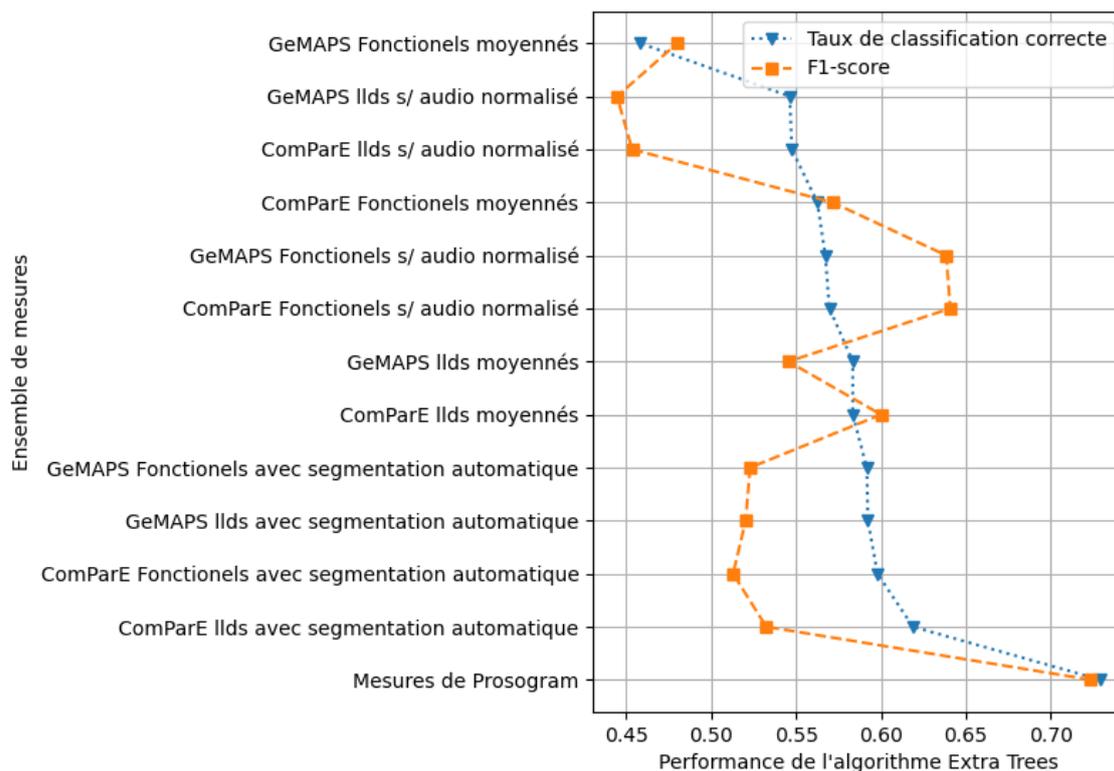


Figure 1. Résultats de l'algorithme Extra Trees sur tous les ensembles de mesures

Après avoir réalisé des analyses sur les ensembles différents, l'algorithme d'apprentissage automatique Extra Trees présentait la meilleure performance globale. Cet algorithme présente des résultats aussi performants que les résultats atteints par Luz et al. (2020) avec des méthodes automatisées de 62,5%. Le résultat le plus intéressant viendrait des mesures extraites automatiquement via Prosogram. Ces mesures ont permis d'atteindre un taux de classification de 72,9%. Ceci indique que ces mesures prosodiques donnent un aperçu pertinent, qui permet une meilleure performance que les ensembles plus complexes.

### Références bibliographiques

- Becker, J. T., Boller, F., Lopez, O. L., Saxton, J., & McGonigle, K. L. (1994). The natural history of Alzheimer's disease. Description of study cohort and accuracy of diagnosis. *Archives of neurology*, 51(6), 585–594. <https://doi.org/10.1001/archneur.1994.00540180063015>
- de la Fuente Garcia, S., Ritchie, C. W., & Luz, S. (2020). Artificial Intelligence, Speech, and Language Processing Approaches to Monitoring Alzheimer's Disease: A Systematic Review. *Journal of Alzheimer's disease : JAD*, 78(4), 1547–1574. <https://doi.org/10.3233/JAD-200888>

- De Roeck, E. E., De Deyn, P. P., Dierckx, E., & Engelborghs, S. (2019). Brief cognitive screening instruments for early detection of Alzheimer's disease: a systematic review. *Alzheimer's research & therapy*, 11(1), 21. <https://doi.org/10.1186/s13195-019-0474-3>
- Eyben, F., Wöllmer, M., & Schuller, B. (2010). Opensmile: The munich versatile and fast open-source audio feature extractor. *Proceedings of the 18th ACM International Conference on Multimedia*, 1459–1462. <https://doi.org/10.1145/1873951.1874246>
- Eyben, F., Scherer, K., Schuller, B., Sundberg, J., André, E., Busso, C., Devillers, L., Epps, J., Laukka, P., Narayanan, S., & Truong, K. (2015). The Geneva Minimalistic Acoustic Parameter Set (GeMAPS) for Voice Research and Affective Computing. *IEEE Transactions on Affective Computing*, 7(2), 190-202, doi: 10.1109/TAFFC.2015.2457417
- Luz, S., Haider, F., Fuente, S.d.l., Fromm, D., & MacWhinney, B. (2020) Alzheimer's Dementia Recognition Through Spontaneous Speech: The ADReSS Challenge. *Proc. Interspeech 2020*, 2172-2176.
- Mertens, P. (2004), Un outil pour la transcription de la prosodie dans les corpus oraux, *Traitement automatique des langues*, 45(2), 109-130.
- Schuller, B., Steidl, S., Batliner, A., Hirschberg, J., Burgoon, J. K., Baird, A., Elkins, A., Zhang, Y., Coutinho, E., & Evanini, K. (2016). The INTERSPEECH 2016 computational paralinguistics challenge: 17th Annual Conference of the International Speech Communication Association. *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH, 08-12-September-2016*, 2001–2005.



# Influence d'un métronome sur la fluence et les gestes de production de parole de personnes qui bégaiant

Maëva GARNIER<sup>1</sup>  
Anneke SLIS<sup>1</sup>  
Clémence DE BELLIS<sup>1</sup>  
Océane BELLENGER<sup>1</sup>  
Christophe SAVARIAUX<sup>1</sup>

<sup>1</sup> Univ. Grenoble Alpes, CNRS, Grenoble INP  
GIPSA-lab, 38000 Grenoble, France

## Introduction

Les personnes qui bégaiant connaissent depuis longtemps l'effet facilitateur de parler avec un métronome (Brady 1969), réduisant la fréquence et la sévérité des disfluences (Hanna & Morris 1977; Kalinowski et al., 2000). Quelques études ont pu montrer que cette amélioration perçue s'accompagnait de modifications objectives de la production de parole au niveau acoustique (Brayton & Conture, 1978 ; Klich & May, 1982 ; Stager, Denman & Ludlow, 1997 ; Davidow, Bothe & Ye, 2011) et aérodynamique (Hutchinson & Navarre, 1977 ; Stager, Denman & Ludlow, 1997).

## Objectifs

L'objectif de cette étude est d'explorer plus en détails les modifications acoustiques, glottiques et articulatoires de la production de consonnes occlusives lorsque la parole est synchronisée avec un métronome, pour des personnes normofluentes ou qui bégaiant, de façon à identifier les aspects de la production qu'un métronome peut contribuer à faciliter.

## Expérience

Seize adultes qui bégaiant (PQB) ont été appariés en âge, genre, et expérience musicale, avec seize adultes normofluents (PNF). Après avoir lu un court texte (READ), ils jouaient à un jeu des 7 différences (GAME), puis répétaient plusieurs phrases en suivant un métronome à 120 BPM (SYNC). Chacune des conditions requerrait la production de mots bi-syllabiques commençant par /pa/ ou /ba/ (En moyenne : 17 occurrences de chaque syllabe par locuteur dans READ et GAME, 21 dans SYNC). Les signaux audio et électroglottographiques de ces productions étaient enregistrés simultanément avec l'activité électromyographique (de surface) de l'orbicularis oris supérieur (OOS) et inférieur (OOI), et avec le mouvement des lèvres (vidéo à 100 images/s).

## Résultats

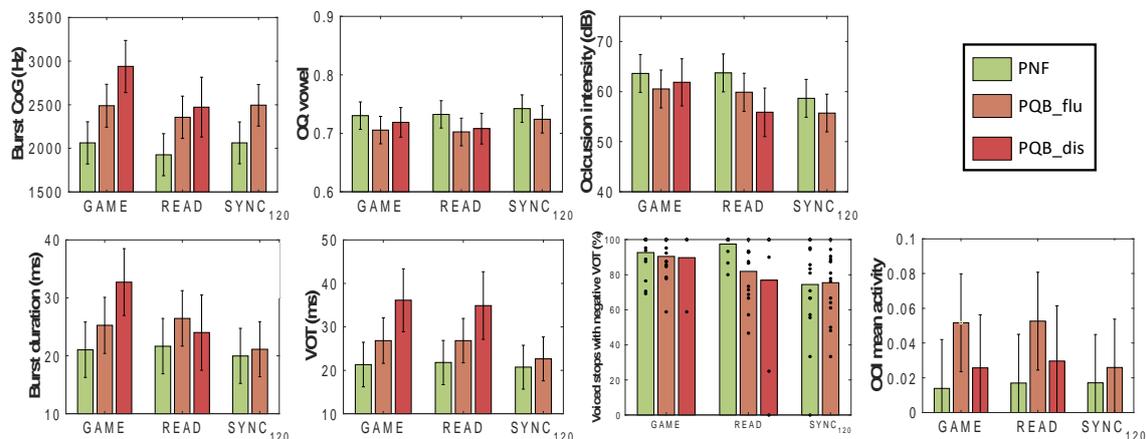
Dans les conditions non rythmiques (GAME, READ), les PQB ont montré des différences significatives au niveau de leurs gestes de production, par rapport à ceux des PNF, non seulement pour leurs syllabes bégayées (PQB\_dis, représentant environ 8% des productions), mais également pour leurs autres syllabes perceptivement fluentes (PQB\_flu).

Ainsi, les PQB ont produit leurs occlusives avec des bruits de plosion plus longs et de barycentre spectral (CoG) plus élevé, des VOT plus longs (pour les /p/), de plus fréquents dévoisements des /b/, des quotients ouverts laryngés (OQ) plus faibles, des prévoisements moins intenses et une plus forte activité de l'OOI (cf. *Figure 1*).

Le métronome a nettement amélioré la fluence des PQB, avec plus aucune disfluence observée sur les syllabes /pa/ ou /ba/ en condition SYNC. En production, plusieurs paramètres, tels que le CoG des bruits ou le OQ, n'ont pas montré de variation significative dans SYNC, comparé aux précédentes conditions GAME et READ, ou ont montré une variation dans le sens du comportement déjà observé chez les PQB dans ces conditions non rythmiques (dévoisements plus fréquents, prévoisements moins intenses). Enfin, plusieurs paramètres ont montré une variation significative dans SYNC, uniquement chez les PQB, de telle façon que leurs productions dans SYNC se « rapprochent » de celles des PNF (raccourcissement du VOT et des bruits de plosion, réduction de l'activité de l'OOI).

### Conclusion

Ces résultats confirment que des différences significatives existent dans les gestes de production des personnes qui bégaièrent, non seulement lors des disfluences, mais également lors de leur parole perceptivement fluente. Parmi ces différences, certaines liées à l'occlusion du conduit vocal et à son relâchement semblent pouvoir être réduites avec l'aide d'un métronome, pouvant expliquer son effet bénéfique sur la fluence.



*Figure 1.* Variations acoustique, glottiques et articulatoires observées entre la parole semi-spontanée (GAME), lue (READ) et synchronisée avec un métronome (SYNC), pour des personnes normofluentes (PNF) ou qui bégaièrent lors des épisodes de disfluence (PQB\_dis) ou lors de leur parole perceptivement fluente (PQB\_flu).

### Références bibliographiques

- Brady, J. P. (1969). Studies on the metronome effect on stuttering. *Behaviour Research and Therapy*, 7(2), 197-204.
- Brayton, E. R. & Conture, E. G. (1978). Effects of noise and rhythmic stimulation on the speech of stutterers. *Journal of Speech and Hearing Research*, 21(2), 285-294

- Davidow, J. H., Bothe, A. K., & Ye, J. (2011). Systematic studies of modified vocalization: speech production changes during a variation of metronomic speech in persons who do and do not stutter. *Journal of fluency disorders*, 36(2), 93-109.
- Hanna, R., & Morris, S. (1977). Stuttering, speech rate, and the metronome effect. *Perceptual and Motor Skills*, 44(2), 452-454.
- Hutchinson, J. M., & Navarre, B. M. (1977). The effect of metronome pacing on selected aerodynamic patterns of stuttered speech: Some preliminary observations and interpretations. *Journal of Fluency Disorders*, 2(3), 189-204.
- Kalinowski, J., Stuart, A., Rastatter, M. P., Snyder, G., & Dayalu, V. (2000). Inducement of fluent speech in persons who stutter via visual choral speech. *Neuroscience letters*, 281(2-3), 198-200.
- Klich, R. J., & May, G. M. (1982). Spectrographic study of vowels in stutterers' fluent speech. *Journal of Speech, Language, and Hearing Research*, 25(3), 364-370.
- Stager, S. V., Denman, D. W., & Ludlow, C. L. (1997). Modifications in aerodynamic variables by persons who stutter under fluency-evoking conditions. *Journal of Speech, Language, and Hearing Research*, 40(4), 832-847.



# Réflexions autour des pauses : quels types de « dysfluencèmes » dans l'aphasie pour quelles mesures en discours continu ?

Halima SAHRAOUI<sup>1</sup>  
Silvia MARTÍNEZ-FERREIRO<sup>2</sup>

<sup>1</sup> Laboratoire de NeuroPsychoLinguistique, EA 4156  
Université de Toulouse 2 Jean Jaurès

<sup>2</sup> Gerontology & Geriatrics Research Group  
Faculty of Health Sciences, University of A Coruña

L'évaluation de la fluence / non fluence en discours spontané chez l'aphasique constitue un enjeu crucial pour une meilleure caractérisation des capacités langagières des patients, d'autant que la fluence peut varier d'une tâche langagière à l'autre et qu'il est essentiel d'identifier les chaînes typiques de dysfluences à des fins diagnostiques et thérapeutiques (ce que nous appelons « dysfluencèmes »). Dans l'aphasie non fluente en particulier, les patients ont recours à des stratégies variées telles que les ellipses ou les auto-corrrections / reformulations pré- ou post-articulatoires (Oomen et al., 2001, Kolk, 2006). Pour mieux caractériser ces stratégies en lien avec la fluidité verbale, nous avons d'abord examiné les aspects temporels et les dysfluences manifestes (répétitions, faux départs, autocorrections et reformulations) chez 5 participants aphasiques (PA) et 6 participants contrôles (PC) à travers deux tâches discursives contrastées : production de discours spontané *vs.* description/récit à partir d'images séquentielles<sup>16</sup>. L'analyse des durées et fréquences des pauses (silencieuses et remplies), du taux articulatoire et des dysfluences met en évidence des différences entre les groupes selon les tâches. Les proportions de dysfluences varient entre tâches pour le groupe PA, avec une augmentation significative en discours d'après images ( $p = 0,016$ ), mais en l'absence de différences significatives concernant les fréquences des pauses entre les deux tâches. Malgré cela, le débit verbal est plus rapide en discours spontané ( $p = 0,054$ ), ce qui traduit une amélioration de la fluidité verbale. Nous avons alors affiné nos analyses suivant les travaux réalisés en discours typique (Grosman et al., 2018 ; Crible et al., 2017, définissant les « fluencèmes »), en dégagant les densités de distribution des pauses, et leur co-occurrence avec des dysfluences ou d'autres structures linguistiques. Nous avons ainsi identifié divers types de « chaînes » composées de pauses associées à certains marqueurs linguistiques ou dysfluences (co-occurrences telles que <pause silencieuse + réparation> ou <marqueur de discours + pause remplie>) qui seraient donc des « fluencèmes » ou « dysfluencèmes » variant en qualité et quantité selon les tâches. Ceux-ci assument un rôle structurel et

---

<sup>16</sup> Les corpus comptent un nombre (N) total de **47.209 mots** : N(PA) = 23.424, N(PC) = 23.785 ; **5.437 pauses silencieuses et remplies** : N(PA) = 4.298, N(PC) = 1.139) ; **2.352 disfluences** : N(PA) = 1.852, N(PC)=500.

fonctionnel d'un point de vue micro- et macro-discursif en discours, et constituent ainsi des indices pertinents afin de réaliser des mesures qui ne se limiteraient pas à une analyse quantitative isolée des pauses, et qui permettent par ailleurs d'envisager les compromis opérés entre fluidité verbale et précision linguistique selon des tâches discursives contrastées dans le cas de l'aphasie.

### Références bibliographiques

- Crible, L., Degand, L. & Gilquin, G. (2017). The clustering of discourse markers and filled pauses: A corpus-based French-English study of (dis)fluency. *Languages in Contrast*, 17, 69-95.
- Grosman, I., Simon, A. & Degand, L. (2018). Variation de la durée des pauses silencieuses : impact de la syntaxe, du style de parole et des disfluences. *Langages*, 211, 13-40.
- Kolk, H. (2006). How language adapts to the brain: An analysis of agrammatic aphasia. In Progovac, L., Paesani, K., Casielles, E., & Barton, E. (Eds.), *The syntax of nonsententials: Multidisciplinary perspectives*, Amsterdam: John Benjamins, 229-258.
- Oomen, C. E., Postma, A. & Kolk, H. H. J. (2001). Pre-articulatory and post-articulatory self-monitoring in Broca's Aphasia. *Cortex*, 37(5), 627-641.

# Index des auteurs

## A

A. N. MacLEOD Andrea, 179  
ABDERRAZEK Sondes, 73  
ADL ZARRABI Aynaz, 57  
AL-TAMIMI Jalal, 127, 187  
AMELOT Angélique, 123  
ANDRE-OBRECHT Régine, 171  
ARIDA Laudy, 169  
AROUN Jeevya, 65  
ASSAL Frédéric, 71, 181  
ASTESANO Corine, 39, 59, 135  
AUCOUTURIER Jean-Julien, 57, 113, 159  
AUDIBERT Nicolas, 13, 85, 99, 151

## B

BAHUAUD Delphine, 49  
BALAGUER Mathieu, 175  
BAQUÉ Lorraine, 3  
BARAKAT Oussama, 119  
BARKAT-DEFRADAS Mélissa, 143  
BASIRAT Anahita, 49  
BEAUMARD Colleen, 147  
BELLENGER Océane, 191  
BENGHANEM Sarah, 159  
BERTICAT Claire, 143  
BOURBON Angéline, 37, 71, 181  
BOURDIN Fabien, 65  
BOURQUI Marion, 35  
BREVET Alexia, 167  
BRIEND Frédéric, 157  
BRUN Lou, 103

## C

CALIMEZ Lorraine, 145  
CHARLIER Brigitte, 23  
CHARUAU Delphine, 68  
CREVIER-BUCHMAN Lise, 13, 37, 85, 123, 151  
CROUZET Olivier, 27

## D

D'ALESSANDRO Daria, 181  
DAVID Céline, 157  
DE BELLIS Clémence, 191  
DEGHORAIN Pauline, 145  
DELAIS-ROUSSARIE Elisabeth, 27  
DELVAUX Véronique, 13, 23, 33, 85, 93, 97, 163  
DEMOLIN Didier, 185  
DÉMOSTHÈNES Isabelle, 81  
DESMETTRES Thibaut, 119  
DIDIRKOVÁ Ivana, 169  
DODANE Christelle, 103, 105  
DUCLOS Laure, 133  
DUEZ Danielle, 167  
DUNIEC Agnieszka, 27

## E

ELMERICH Amélie, 123  
EMBARKI Mohamed, 119

## F

FAGNIART Sophie, 13, 23, 85  
FARINAS Jérôme, 59, 129, 175  
FERRÉ Sandrine, 51, 157  
FERY Patrick, 145  
FONTAN Lionel, 43  
FOUGERON Cécile, 37, 71, 99, 181  
FREDOUILLE Corinne, 73, 133, 167  
FUMEL Jules, 49  
FUSAROLI Riccardo, 49

## G

GARNIER Maëva, 53, 191  
GASNIER Marlène, 51  
GAVARET Martine, 159  
GHIO Alain, 73, 167  
GINDRE Anne-Flore, 135  
GOESEELS Eva, 163  
GRANDON Bénédicte, 17, 51

GRAVELLIER Lila, 129  
GUEROUAOU Nadia, 113  
GUERRY Nathalie, 145  
GUHA Rudradeep, 159  
GUITARD-IVENT Fanny, 39

## H

HAMON Sophie, 133  
HARMEGNIES Bernard, 13, 23, 33, 85,  
97  
HENNECHART Manon, 65  
HIRSCH Fabrice, 103, 105, 169  
HUBERLANT Anne, 23  
HUET Kathy, 13, 23, 33, 85, 93, 97, 163

## J

JAYR Manon, 167  
JORDAN Eric, 187  
JOUCLA Coralie, 159  
JUCLA Mélanie, 39  
JUDENNE Richard, 171

## K

KODRASI Ina, 71

## L

LAGANARO Marina, 35, 71, 181  
LALAIN Muriel, 73, 133, 167  
LANCHEROS Monica, 35  
LATINUS Marianne, 157  
LAZARD Diane, 65  
LE COZ Maxime, 129  
LEVEQUE Nathalie, 71  
LÆVENBRUCK Hélène, 19, 155

## M

MABBOUX Elisa, 133  
MACHART Laura, 19, 155  
MAEDA Shinji, 123  
MALVY Joëlle, 157  
MARTEL-SAUVAGEAU Vincent, 7  
MARTÍNEZ-FERREIRO Silvia, 43, 195  
MARX Tania, 119  
MAUBLANC Paul, 113  
MAUCLAIR Julie, 75, 109, 171

MELONI Geneviève, 179  
MÉNARD Lucie, 19, 81  
MEUNIER Christine, 73  
MOOTASSIM-BILLAH Sofiana, 89

## N

NACCACHE Lionel, 57  
NEVEU Fabrice, 129  
NOPERE Martin, 97  
NOWAKOWSKA Aleksandra, 43

## P

P. CIRCIU Marta, 185  
P. MARTIN Vincent, 147  
PAPON Jean-François, 123  
PARIENTE Jérémie, 39  
PERNON Michaela, 71, 181  
PERRIER Pascal, 53  
PETTIROSSI Amelia, 13, 85, 151  
PHILIP Pierre, 147  
PICCALUGA Myriam, 13, 23, 33, 85, 93,  
163  
PILLOT-LOISEAU Claire, 13, 65, 85  
PINQUIER Julien, 75, 109, 129, 175  
POMMEE Timothy, 171  
PORTET François, 117  
POUCHOULIN Gilles, 167  
PRINCE Typhanie, 43  
PRUVOST-ROBIEUX Estelle, 159

## Q

QUINTAS Sebastião, 75, 109

## R

RABARISOA Faraniaina, 133  
RAYMOND Michel, 143  
REMMAN Reina, 13, 85  
RINGEVAL Fabien, 117  
ROBERT Stephan, 119  
ROLAND Virginie, 13, 33, 85, 93, 97,  
163  
RONDELLO David, 133  
ROQUES Cwiosna, 105  
ROUAS Jean-Luc, 147  
ROUCHEZ SLIMANI Amira, 13

RUIGENDIJK Esther, 17

## S

SAHRAOUI Halima, 43, 135, 195

SAVARIAUX Christophe, 53, 191

SCHOENTGEN Jean, 89

SILLERESI Silvia, 157

SITBON Laurianne, 5

SLIS Anneke, 53, 191

SOCK Rudolph, 68

## T

TERNY Lola, 13, 85, 93

TIEDE Mark, 19

TRESALLET Christophe, 65

## V

VAIVA Guillaume, 113

VAN BOGAERT Lucie, 155

VAN DER MAREN Anne-Sophie, 145

VAN GESTEL Dirk, 89

VAUCHEREY Coralie, 119

VAXELAIRE Béatrice, 68

VAYSSE Robin, 59

VERHAEGEN Clémence, 13, 85, 93,  
163

VIALLET François, 167

VILAIN Anne, 19, 155

VILLAIN Marie, 57

## W

WEED Ethan, 49

WEI Dongjun, 119

WESTERMEYER Laura, 13

WINLING Noémie, 179

WOISARD Virginie, 73, 167, 175

WU Yaru, 147

## Y

YANG Lulin, 143

## Z

ZHOU Yongxin, 117







ISBN : 978-2-917490-35-8