GLOSSA

Revue Scientifique en Orthophonie Logopédie

Numéro 142, 2024

Directrice de publication :
Sylvia Topouzkhanian

Équipe de rédaction :
Rédactrice en chef :
Agnès Witko
agnes.witko@univ-lyon1.fr

Directrice de production et administratrice web :
Lydie Batilly-Gonin

Secrétaires de rédaction :
Florence Baldy-Moulinier
Bénédicte Bordet-Boullet
Marianne Le Floch-Bazin
Mathilde Oudry

e-ISSN : 2117-7155

Site web :
https://www.glossa.fr

Licence :
Ce travail est disponible sous licence Creative Commons Attribution 4.0 International.

Editorial Glossa 142.
AGNES WITKO & ETIENNE SICARD

Tâche pour l'évaluation de la production et la compréhension de syntagmes nominaux complexes impliquant l'accord en genre en français.
French assessment task for production and comprehension of complex noun phrases and gender agreement.
PHAEDRA ROYLE & NATACHA TRUDEAU

Évaluation des dysarthries et des apraxies de la parole : quelles approches ?
Assessment approaches for dysarthria and apraxia of speech.
MICHAËLA PERNON

Méthode d'analyse des pauses pour l'évaluation de troubles cognitifs subtils appliquée au Cancer-Related Cognitive Impairment.
Method for the analysis of pauses for assessing subtle cognitive impairment applied to Cancer-Related Cognitive Impairment.
AMÉLIE RICHARD, FABRICE HIRSCH & SOPHIE JACQUIN-COURTOIS

Editorial Glossa 142.

Auteurs :
Agnès Witko[1,2]
Etienne Sicard[3]

Affiliations :
[1] UCBL - Laboratoire DDL, Lyon,France
[2] Rédactrice en chef de Glossa
[3] INSA Toulouse, Toulouse, France.

Autrice de correspondance :
Agnès Witko
agnes.witko@univ-lyon1.fr

Comment citer cet article :
Witko, A., Sicard, E. (2024). Editorial Glossa 142. *Glossa*, 142, 2-6. https://doi.org/10.61989/1fj2bm45

e-ISSN : 2117-7155

Licence :
© Copyright Agnès Witko, Etienne Sicard, 2024
Ce travail est disponible sous licence Creative Commons Attribution 4.0 International.

L'ORTHOPHONIE AU-DELÀ DES CONSULTATIONS…

Agnès WITKO, Rédactrice en chef

Vous vous intéressez à l'orthophonie au-delà des consultations ? Vous cherchez à actualiser vos connaissances en permanence ? Vous souhaitez développer vos compétences cliniques en vous basant sur des sources scientifiques ? Vous préférez la science à l'idéologie, le progrès à la nostalgie ? Et tout cela, pour mieux appréhender la complexité des structures et des processus du langage, de la communication et de la parole ?

La revue Glossa propose des études de plus en plus détaillées, de plus en plus sophistiquées, au service de la compréhension des difficultés parfois subtiles et invisibles, parfois envahissantes et accablantes, que rencontrent les patients qui consultent en orthophonie.

A l'heure de la science ouverte, Glossa porte son intérêt sur les données collectées en recherche, sur les soins du langage et de la communication : une liste de mots, une répétition d'énoncé, un temps de traitement, etc. Les données brutes suscitent un intérêt à juste titre, dans la mesure où ces dernières renvoient pour partie au soin courant et à la matière brute que les cliniciens et cliniciennes évaluent, analysent et interprètent.

Parce que ces compétences cliniques peuvent nourrir la connaissance, les ponts entre clinique et théorie (et l'inverse) se multiplient dans différents secteurs, pour exemples, dans la précision des diagnostics reliés à la sévérité des troubles, dans l'explication des décisions de soins en lien avec les besoins des patients, dans le développement d'outils fiables proposant des mesures précises et adaptées en soin courant. Devant cette richesse et cette complexité, les données collectées passent par des méthodologies qui jouent la carte de l'interconnexion.

Nous verrons dans la deuxième partie de cet éditorial avec Etienne Sicard, comment l'intelligence artificielle et ses différents modèles pourraient offrir des moyens à cette nouvelle gestion de la complexité, dans les sciences biomédicales, humaines et sociales.

Revenons aux trois articles de ce dernier numéro de l'année 2024 qui explorent des facteurs au cœur du réacteur du langage, puisqu'il est question de syntaxe et de production de parole.

Tâche pour l'évaluation de la production et la compréhension de syntagmes nominaux complexes impliquant l'accord en genre en français,

par Phaedra Royle et Natacha Trudeau

Dans leur étude, Phaedra Royle et Natacha Trudeau présentent un outil composé de quatre casse-têtes ludiques de difficulté croissante, avec des images d'objets et d'animaux de couleur et de taille variées, proposées à 190 enfants francophones âgés de 3 à 9 ans. L'objectif consiste à évaluer rapidement les compétences en production de syntagmes nominaux avec des adjectifs de grandeur ou de couleur, et l'accord en genre du syntagme nominal en français.

Les résultats permettent de situer la performance d'un enfant par rapport à ses pairs, de multiples façons. Les seuils de risque, les taux de réussite et les erreurs typiques à chaque âge sont présentés, l'âge auquel les différentes tâches devraient être réussies par la majorité des jeunes est répertorié. Les erreurs morphologiques ou syntaxiques, typiques ou non des enfants grandissant dans des milieux francophones sont aussi rapportées. Ces points de repères quantitatifs et qualitatifs pourront alimenter l'évaluation clinique des orthophonistes, en pointant des domaines de difficulté identifiés dans la morphosyntaxe. Cette nouvelle tâche est tout à fait complémentaire des outils d'évaluation plus classiques pour évaluer les compétences grammaticales, ou même de l'analyse de corpus spontané.

Evaluation des dysarthries et des apraxies de la parole : quelles approches ?

par Michaëla Pernon

Par une comparaison d'approches perceptive, acoustique et cognitive, Michaëla Pernon s'intéresse à l'évaluation des troubles moteurs acquis de la parole (Motor Speech Disorders, MSDs), dysarthrie et/ou apraxie de la parole (Apraxia of Speech, AoS). Trois études ont été menées chez quatre groupes de locuteurs : neurotypiques, ou pathologiques présentant soit une AoS, soit une dysarthrie hypokinétique, ou encore une dysarthrie mixte flasque-spastique dans la sclérose latérale amyotrophique. Une partie des tâches d'évaluation, basées sur le corpus MonPaGe-2.0.s, a porté sur une classification auditivo-perceptive multijuges des MSDs. De plus, une analyse phonético-acoustique discrète a concerné des locuteurs aux taux de précision les plus bas et les plus élevés, un paradigme de double tâche évaluant, chez des locuteurs neurotypiques, les effets d'une demande attentionnelle et exécutive sur des tâches de parole et non verbales. Sur un plan clinique, les résultats montrent que les approches d'évaluation perceptive et acoustique ne permettent pas de diagnostiquer les MSDs très légers, mais produisent une caractérisation quantifiée des profils des MSDs sur trois dimensions de parole : la nature de la tâche de parole (degré de familiarité, degré d'automaticité, etc.), la complexité articulatoire du corpus mis en jeu par des variables psycholinguistiques/phonétiques et les conditions de production de la tâche de parole. En complément, l'examen des propriétés cognitivo-motrices de parole selon le type de tâche, offre aux orthophonistes la possibilité d'affiner le diagnostic différentiel des MSDs. Dans l'attente des systèmes de reconnaissance automatique de la parole, le recours à un raisonnement guidé tel que proposé dans cette étude permettra d'affiner le diagnostic différentiel de MSDs.

Méthode d'analyse des pauses pour l'évaluation de troubles cognitifs subtils appliquée au Cancer-Related Cognitive Impairment,

par Amélie Richard, Fabrice Hirsch et Sophie Jacquin-Courtois

Le manque d'outils d'évaluation suffisamment sensibles pour déceler certains Troubles NeuroCognitifs mineurs (TNCm) a conduit Amélie Richard et ses coauteurs à rechercher des marqueurs comportementaux auprès de patients présentant un trouble cognitif induit par un cancer et ses traitements, ou Cancer-Related Cognitive Impairment (CRCI). Caractérisées principalement par une plainte mnésique et/ou linguistique, la fréquence des CRCI et leur intensité ne sont pas corrélées aux scores des tests orthophoniques ou neuropsychologiques. Une analyse fine de la parole a été conduite auprès de treize participantes post-cancer du sein et treize participantes contrôles saines qui ont raconté une histoire à partir d'une séquence d'images. Les pauses silencieuses, les pauses remplies et les allongements vocaliques ont fait l'objet d'une analyse statistique montrant que seule la durée des pauses silencieuses des participantes post-cancer du sein était significativement plus longue

que celle des contrôles. Cette étude décrit en détails les étapes d'analyse des pauses en vue de son applicabilité clinique, visant à répondre à une plainte de patientes rapportant un changement cognitif pour lequel l'autonomie est préservée, et qui peuvent, le cas échéant, porter un trouble invisible qui réduit leur qualité de vie.

Qu'il s'agisse de syntaxe ou de parole, les marqueurs de difficulté ou de trouble demandent des analyses très fines que la recherche nous révèle. Les compétences des orthophonistes passeront par ces nouvelles approches, intégrées progressivement à leur expertise, notamment grâce aux outils numériques. Parmi ceux-ci, Etienne Sicard nous présente comment l'intelligence artificielle peut établir des liens possibles avec l'orthophonie, partant du principe que tous les moyens sont à explorer pour communiquer et traduire la science, de manière à soutenir les praticiens et praticiennes de demain.

VERS UNE JUMELLE NUMÉRIQUE DE L'ORTHOPHONISTE ?

Etienne SICARD, Professeur, INSA Toulouse, directeur de recherches au LURCO/ UNADREO.

Le moins que l'on puisse dire, c'est que l'intelligence artificielle (IA) fait débat. Dans l'éditorial du N°141 de Glossa[1], Natacha Trudeau discutait récemment de l'apport de l'IA, de la réalité virtuelle (VR) et des plates-formes d'apprentissage interactives telles que la « clinique virtuelle en orthophonie et audiologie »[2] pour l'aide à l'acquisition des compétences des futurs orthophonistes.

Le règlement européen sur l'intelligence artificielle (IA)[3], appelé « IA Act » est entré en vigueur en 2024. Il vise à favoriser le développement et déploiement responsable de l'intelligence artificielle dans l'Union Européenne (UE), tout en précisant les exigences et obligations pour les développeurs et fournisseurs de solutions. L'IA Act considère notamment que les logiciels médicaux fondés sur l'IA doivent respecter des exigences strictes en matière de risques pour la santé, de qualité du diagnostic et de prise en soin, ou encore de fourniture et gestion d'informations au patient.

Dans le contexte de l'aide à la prise en soin orthophonique, on peut considérer plusieurs niveaux d'autonomie de l'IA, le niveau le plus haut pouvant être considéré comme la jumelle numérique[4] de l'orthophoniste (JUNO). Nous pouvons distinguer 5 niveaux d'autonomie et d'intelligence, un peu à la manière des standards de conduite automatique de la voiture autonome[5].

Le niveau zéro consiste à se passer de tout outil numérique dans la prise en soin, l'une des justifications les plus communes étant la surexposition aux écrans des patients, notamment leur addiction aux outils numériques, parmi lesquels se trouve en première place le smartphone. Cette approche laisse une grande place à l'interprétation, à la subjectivité, au jugement et donc au risque d'erreur. Environ 1/3 des orthophonistes restent réfractaires aux outils numériques lors du bilan, mais seulement 10% en prise en soin, selon une étude réalisée auprès de 200 Speech-Language Therapists (SLT) aux USA[6].

Le premier niveau de numérisation, qui intègre les précurseurs de l'IA, permet la classification objective de l'altération des performances du patient, grâce à des indicateurs dont la fiabilité a été mesurée, en général par comparaison avec une réalité terrain. Ces approches reposent sur un protocole strict, de bonnes conditions de passation et des orthophonistes bien formées. Dans le domaine de la voix, les indicateurs d'altération du /a:/ tenu de l'outil VOCALAB font partie de cette catégorie[7]. Dans celui de la parole, les indicateurs d'altération de séries de /Pataka/ font aussi appel à des notions de segmentation automatisée, d'évaluation des caractéristiques des voyelles et consonnes, puis de classification entre

1 Glossa Revue Scientifique en Orthophonie Logopédie Glossa N°141, pp. 2-5

2 https://laborenato.ca/clinique-artificielle/

3 https://www.vie-publique.fr/questions-reponses/292157-intelligence-artificielle-le-cadre-juridique-europeen-en-6-questions

4 https://fr.wikipedia.org/wiki/Jumeau_num%C3%A9rique

5 https://www.sae.org/blog/sae-j3016-update

6 Albudoor, N., (2021). Factors influencing US speech and language therapists' use of technology for clinical practice. International Journal of Language & Communication Disorders, 56(3), 567-582.

7 Sicard, E., Menin-Sicard, A., Perriere, S. (2013) Développement et validation d'outils de mesures de la qualité de la voix dans le logiciel VOCALAB, Glossa 113, pp. 63-80

fluence normale, altérée ou pathologique[8]. D'une certaine façon, le protocole MonPaGe utilisé par Michaëla Pernon dans son article sur l'évaluation des dysarthries et des apraxies de la parole publié dans ce numéro, entre aussi dans cette catégorie d'outils.

Un niveau supplémentaire (niveau 2) consiste à appliquer des techniques de « Machine Learning » sur de grands corpus de données patients pour permettre aux orthophonistes de disposer d'un avis étayé sur la gravité de certains troubles. On peut citer dans le domaine de la parole les travaux visant à classifier les dysarthries[9]. A notre connaissance, il n'existe pas encore de tels outils déployés en France à destination des orthophonistes. Nous avons pour notre part développé un outil prototype de classification des syndromes Parkinsoniens, mais avons reculé devant la responsabilité sociétale écrasante que ce type d'analyse représente, et l'incapacité à atteindre le très haut niveau de robustesse requis pour une aide au diagnostic précoce. Il nous semble que seule une approche concertée entre spécialistes internationaux, avec une vision à long terme, une très haute exigence de qualité et d'éthique et un partenariat de confiance recherche/industrie, permettrait d'atteindre ce palier. Les grands programmes cadres européens sont une piste possible, mais les financements sont en général assurés sur des périodes courtes, avec des exigences et critères d'éligibilité qui rendent la marche souvent insurmontable, notamment pour des laboratoires de recherche orthophonique tels que le LURCO[10] et le SURO[11].

Le niveau 3 consiste à donner une réelle autonomie à la machine pour donner des conseils éclairés à l'orthophoniste, au moins sur certains domaines relativement bien structurés, bornés et se prêtant favorablement à une aide de la machine. Cela suppose notamment une bonne connaissance des méthodologies de bilan et des approches de prise en soin, notamment en lien avec les pratiques probantes[12], et la mise à disposition de grands corpus de données patients. Qui dit conseils éclairés pose immédiatement la question de l'éthique de l'IA. Une mise en scène sous forme de procès de l'IA a été réalisée en 2024 par le Centre National de la Recherche Scientifique (CNRS)[13], avec des chefs d'accusation : « peu explicable, porteuse de biais, facile à manipuler, aggravant les inégalités sociales, déshumanisant les échanges ou encore menaçant nos démocraties, etc. », et la réponse de différents experts du domaine.

Nous avons démarré en 2023 le projet MADLEN[14] qui vise à donner des conseils et recommandations bienveillantes dans le domaine de la voix et de la parole, qui s'approche donc, au moins dans l'esprit, de ce niveau 3 d'une jumelle numérique. Nous nous sommes basés sur plusieurs milliers de cas de voix et parole normales, altérées ou pathologiques, ainsi que 100 études de cas de la part d'orthophonistes[15]. Une des propositions clé consisterait à conseiller des exercices pertinents en lien avec les performances du patient, sur la base d'une matrice de décision exploitant une vingtaine d'indicateurs d'altération de la voix et de la parole. Un prototype est en cours de mise au point, notamment en lien avec l'AFAF (Association Française de l'Ataxie de Friedreich) et devrait passer par de longs mois de tests auprès d'orthophonistes en vue d'évaluer la pertinence de l'approche et la qualité du service rendu.

De manière plus futuriste, on peut imaginer un niveau 4 de virtualisation qui reposerait sur l'analyse d'un nombre très important d'interactions patient/thérapeute, grâce à l'analyse vidéo, la reconnaissance vocale, l'analyse lexicale et l'annotation clinique, permettrait d'alimenter un outil d'interaction comme un ChatBot, une IA générative ou un clone adapté au contexte de la prise en soin orthophonique. On peut supposer qu'il faut y associer une exploitation massive de la science ouverte, des méta-analyses sur les différentes approches thérapeutiques

8 Sicard, E., & Menin-Sicard, A. (2020). La diadococinésie et son application en orthophonique clinique. https://hal.science/hal-02512856/

9 Al-Ali, A., (2024). The Detection of Dysarthria Severity Levels Using AI Models: A Review. IEEE Access.

10 https://www.unadreo.org/lurco/presentation-du-lurco/

11 https://surorthophonie.com/

12 Marie-Anne Schelstraete, Pour un meilleur dialogue entre la recherche et la clinique, Glossa N°136, Editorial

13 https://youtube.com/@CNRS-FIDLE

14 https://www.formationsvoixparole.fr/labo/madlen/

15 Sicard, E., (2024). Etude de cas de pathologies de la parole dans le cadre de la prise en charge orthophonique – Volume III, https://hal.science/hal-04497896

mises en œuvre, pour augmenter la pertinence de cet outil. Cela implique aussi une capacité à mesurer, modéliser et restituer les caractéristiques de pathologies très diverses, une sorte de généralisation numérique d'ouvrages tels que la classification des pathologies de la voix de Verdolini[16] ou les études de cas de prise en soin de la parole de Stemple[17].

Enfin, le dernier niveau serait une véritable jumelle numérique de l'orthophoniste, qui en plus de l'intelligence artificielle, intègrerait des mécanismes d'apprentissage automatique et d'analyse au long cours des données issues des bilans et des prises en soin, pour affiner ses modèles de simulation en permanence, augmenter en pertinence, tout en restant sous le contrôle scientifique et éthique d'un « comité des sages », probablement abrité par une société savante telle que l'UNADREO[18]. A ce niveau de sophistication, on peut penser à une implémentation sous forme de robot humanoïde tel que le Tesla Bot[19], d'autant que les investissements en recherche et développement semblent hors d'atteinte des circuits de financements classiques et risquent de ne pouvoir être supportés que par de très grandes entreprises innovantes, voire disruptives.

Comme à chaque fois lors de l'apparition d'une technologie de rupture, nous pouvons être partagés entre enthousiasme et crainte. Enthousiasme vis-à-vis d'une capacité à fournir aux cliniciens des informations à haute valeur ajoutée, basée sur des approches probantes étayées par la littérature scientifique, tout en étant ciblées au mieux des besoins des patients ; enthousiasme à l'idée d'accroître l'expertise des orthophonistes et les accompagner dans l'évaluation des progrès du patient, avec des processus de décision ouverts et transparents.

Mais aussi la crainte des conséquences d'une collecte de données sensibles de milliers de patients, des risques liés au stockage d'informations incluant des données sur la pathologie, l'évolution des performances des patients, la prescription des exercices, etc. ; crainte d'une incapacité à garantir que les fournisseurs de solutions à base d'IA offrent la transparence nécessaire et un accès compréhensible à leurs algorithmes de décision, qu'ils ne manipulent pas les données et ne portent pas préjudice aux patients ; crainte enfin qu'en prenant en charge des tâches de plus en plus complexes, ces jumelles numériques induisent un mouvement de déresponsabilisation des cliniciens, en leur faisant perdre progressivement leur libre-arbitre, voire en se substituant peu à peu à eux.

Les sciences orthophoniques, naturellement au cœur des sciences biomédicales et des sciences humaines et sociales, doivent dès à présent composer avec l'IA générative et peut-être dans un futur proche avec les jumelles numériques de l'orthophonie. La profession est partagée entre la volonté de conserver des pratiques traditionnelles considérées comme raisonnables et sûres et l'intuition que des approches innovantes comme l'IA, la réalité virtuelle, voire la jumelle numérique, vont révolutionner leur pratique. Encore peu alimentée aujourd'hui sur le sujet, la revue GLOSSA pourrait être un vecteur de diffusion du savoir sur ce sujet complexe, à la fois intrigant et fascinant, en mettant en avant ses applications pratiques pour la clinique orthophonique.

16 Verdolini, K., (2014). Classification manual for voice disorders-I. Psychology Press.

17 Stemple, J. C., & Hapner, E. R. (2019). Voice therapy: clinical case studies. Plural Publishing.

18 https://www.unadreo.org/

19 https://en.wikipedia.org/wiki/Optimus_(robot)

Tâche pour l'évaluation de la production et la compréhension de syntagmes nominaux complexes impliquant l'accord en genre en français.

Autrices :
Phaedra Royle[1, 2, 3]
Natacha Trudeau[1, 4, 5]

Affiliations :
[1]École d'orthophonie et d'audiologie, Université de Montréal, Canada.
[2]Centre for Research on Language, Brain and Music (CRBLM), Montréal, Canada.
[3]Centre de recherche sur le cerveau et l'apprentissage (CIRCA), Université de Montréal, Canada.
[4]CRIR–Institut universitaire sur la réadaptation en déficience physique de Montréal (IURDPM), Canada.
[5]Institut universitaire de réadaptation en déficience physique de l'université de Montréal (IURDPM), Canada.

Autrice de correspondance :
Phaedra Royle
phaedra.royle@umontreal.ca

Dates :
Soumission : 30/03/2024
Acceptation : 30/08/2024
Publication : 28/12/2024

Comment citer cet article :
Royle, P. & Trudeau, N. (2024). Tâche pour l'évaluation de la production et la compréhension de syntagmes nominaux complexes impliquant l'accord en genre en français. *Glossa*, 142, 7-28. https://doi.org/10.61989/z4dt9f61

e-ISSN : 2117-7155

Licence :
© Copyright Phaedra Royle, Natacha Trudeau, 2024
Ce travail est disponible sous licence Creative Commons Attribution 4.0 International.

Contexte. Peu d'outils normés existent pour l'évaluation du vocabulaire d'adjectifs de couleur et de grandeur, de la syntaxe et de l'accord en genre du syntagme nominal en français.

Objectif. Cette étude présente un outil, composé de quatre casse-têtes ludiques, qui peut évaluer rapidement les compétences en production de syntagmes nominaux avec des adjectifs de grandeur ou de couleur.

Hypothèses. Nous observerons une consolidation de l'accord de l'adjectif entre les âges de 3 à 6 ans, mais une maîtrise de l'assignation du genre chez les plus jeunes.

Méthode. Les données de 190 enfants francophones âgés de 3 à 9 ans sont présentées. Nous rapportons des résultats sur quatre tâches de difficulté croissante impliquant des adjectifs de couleur et de grandeur (p. ex. le petit canard vert).

Résultats. Les résultats permettent de situer la performance d'un enfant par rapport à ses pairs, de multiples façons. Les seuils de risque, les taux de réussite et les erreurs typiques à chaque âge sont présentés. De plus, l'âge auquel les différentes tâches devraient être réussies par la majorité des jeunes est répertorié. Les erreurs morphologiques ou syntaxiques typiques (ou non) des enfants grandissant dans des milieux francophones sont aussi rapportées.

Conclusion. Ces points de repères quantitatifs et qualitatifs pourront alimenter l'évaluation clinique des orthophonistes, que ce soit avec la tâche ou à partir de corpus de langage, et éclaireront le raisonnement menant au diagnostic clinique en contribuant à l'ensemble des données prises en compte dans un processus de raisonnement clinique.

Mots-clés : accord grammatical ; syntagme nominal ; morphosyntaxe ; production induite ; adjectifs ; genre grammatical.

French assessment task for production and comprehension of complex noun phrases and gender agreement.

Context. Few standardized tools exist to assess vocabulary, syntax and grammar in French.

Objective. This study presents a tool, made of four entertaining puzzles, that can quickly assess proficiency in the production of noun phrases with color or size adjectives. Hypothesis: We'll observe a consolidation of adjective agreement between the ages of 3 and 6, but a mastery of gender assignment in the youngest children.

Method. Data from 190 French-speaking children aged 3 to 9 are presented. We report results on four tasks of increasing difficulty involving adjectives of color and size (e.g., le petit canard vert 'the small green duck').

Results. Our results enable one to situate a child's performance in relation to his or her peers, in a variety of ways. Risk thresholds, success rates and typical errors for each one-year age range are presented, as well as at what ages the tasks should be passed by a majority of children. Morphological or syntactic errors typical (or not) of children growing up in French-speaking environments are also reported.

Conclusion. These quantitative and descriptive benchmarks can inform clinical assessment of speech-language therapists, whether with the task or from language corpora, and will inform reasoning leading to clinical decisions.

Key words: grammatical agreement; noun phrase; morphosyntax; elicited production; adjectives; grammatical gender.

INTRODUCTION

Les orthophonistes du Québec et du Canada œuvrant auprès des enfants francophones doivent souvent évaluer leurs habiletés de communication et de langage sans avoir recours à des tâches élaborées spécifiquement pour le français parce que des tâches adaptées au français laurentien (parlé au Québec et en Ontario, au Canada) n'existent pas ou ne ciblent pas les éléments recherchés, par exemple l'accord en genre. Par conséquent, les outils disponibles ne font souvent pas référence à des éléments clés de la structure grammaticale du français en général mais particulièrement de sa variante laurentienne. Au Québec 9,4% des enfants de 5 ans ont un trouble développemental du langage (TDL) et 72% d'entre eux présenteront toujours des difficultés à l'âge de 12 ans (Ordre des orthophonistes et audiologistes du Québec, OOAQ, 2014). Le ministère de la santé et des services sociaux (MSSS) constate depuis vingt ans dans sa politique familiale la nécessité d'assurer auprès des jeunes enfants une stimulation précoce des habiletés cognitives et langagières (OOAQ, 2014).

Le développement d'outils adaptés à la langue française afin d'évaluer le développement du langage des jeunes enfants (par exemple dans une optique clinique) ainsi que les habilités linguistiques particulières des enfants d'âge pré-scolaire (par exemple en recherche) a été entamé par plusieurs chercheurs au Québec et au Canada (Filiatrault-Veilleux et al., 2016 ; Paul & Thordardottir, 2019 ; Pesco & O'Neill, 2016 ; Thordardottir, 2005; Thordardottir & Namazi, 2007 ; Thordardottir et al., 2010 ; Trudeau et al., 1999 ; Trudeau & Sutton, 2011). Avant l'an 2000, les tests adaptés au français laurentien étaient rares et surtout non standardisés (Bouchard et al., 2009), mais depuis peu, cet état de fait change.

Les premiers travaux ont voulu caractériser le développement du langage de l'enfant francophone parlant le français laurentien de moins de trois ans par des inventaires de mots et gestes (Trudeau et al., 1999), et chez l'enfant de moins de cinq ans par l'analyse de corpus de langage spontané (Thordardottir, 2005 ; Thordardottir & Namazi, 2007). Ces approches permettent un point de vue global sur le développement du langage de l'enfant. On sait cependant que chaque langue présente des éléments plus difficiles que d'autres à maîtriser, peu importe la présence ou non d'un TDL. Conséquemment, une appréciation globale du langage de l'enfant peut ne pas relever des difficultés subtiles ou spécifiques à certaines sous-composantes du langage. Le français présente de nombreux domaines de difficulté identifiés dans la morphosyntaxe, dont certains sont particulièrement en jeu durant la période préscolaire : 1. la flexion verbale (Jakubowicz, 2003 ; Jakubowicz & Nash, 2001 ; Rose & Royle, 1999 ; Royle, 2007 ; Royle & Thordardottir, 2008), 2. les pronoms clitiques (Grüter, 2005 ; Jakubowicz et al., 1998) et 3. l'accord en genre (Jakubowicz, et al., 1998 ; Jakubowicz & Roulet, 2007 ; Roulet-Amiot & Jakubowicz, 2006 ; Royle & Reising 2019 ; Royle & Stine, 2013 ; Royle et al., 2010). Ce dernier domaine nous intéresse plus particulièrement car non seulement il présente un défi à l'apprentissage du français du moins en bas âge (Roulet-Amiot & Jakubowicz, 2006 ; Royle & Valois, 2010), mais n'est pas présent en anglais, et ne peut donc pas faire partie de tâches adaptées de l'anglais.

Nous présentons ici un outil d'évaluation qui pourra servir de premier jalon dans l'identification de troubles au niveau du développement de l'accord du genre en français mais aussi de la syntaxe du syntagme nominal (aussi appelé *groupe nominal*). Cet outil pourra aussi servir à la vérification de la maîtrise ou de la connaissance de concepts de couleur et de grandeur ainsi qu'au développement de la référence, en ce qui a trait à l'utilisation des formes définies et indéfinies des déterminants (articles) dans le syntagme nominal. Cette tâche est complémentaire aux outils d'évaluation plus généraux, tels le CELF-5 pour les compétences grammaticales (Wiig et al., 2019), et l'analyse de corpus spontané, car elle permet de vérifier un domaine spécifique du développement de la morphosyntaxe du français en utilisant un outil adapté aux capacités linguistiques et pragmatiques de très jeunes enfants âgés de 3 à 6 ans. Nous espérons que les spécialistes concernés pourront y voir un outil utile au dépistage précoce d'un trouble de la communication ainsi qu'au dépistage précoce du risque qu'un tel trouble émerge.

Particularités de la structure du syntagme nominal du français en particulier sur l'accord intranominal

Le syntagme nominal s'articule autour du nom ou pronom et, dans la plupart des cas en français, d'un déterminant. Le syntagme peut aussi contenir un ou plusieurs adjectifs qualifiant le nom. Cette structure est acquise très tôt, dès les premiers

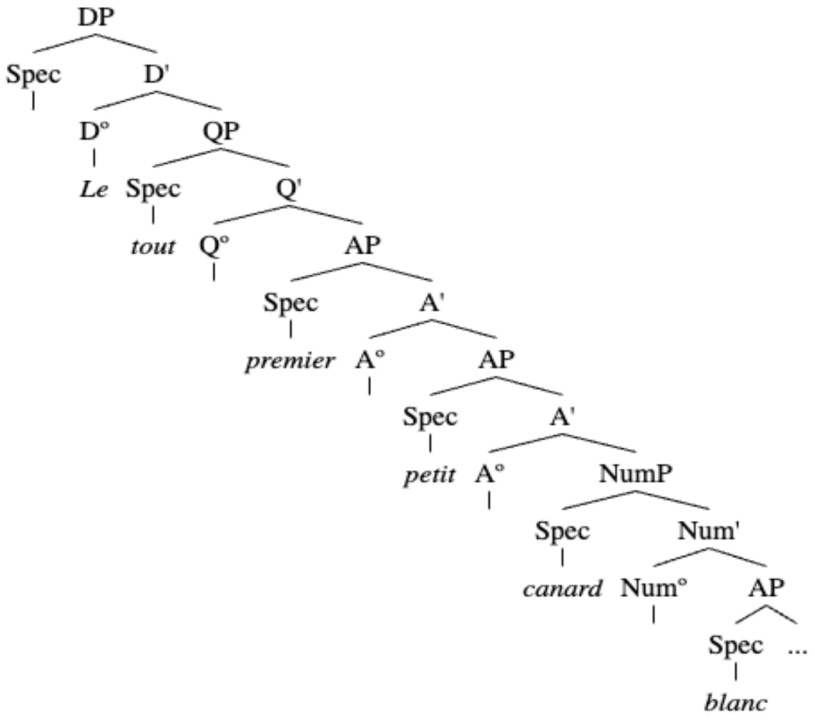


FIGURE 1 : La structure syntaxique du syntagme nominal en français (adaptée de Ronat, 1977)

énoncés (voir p. ex. Valois & Royle, 2009 ; Valois et al., 2009) et le lexique des jeunes francophones contient une majorité de noms (Trudeau & Sutton, 2011). Dès les premières combinaisons de mots, on peut déjà observer des structures de type déterminant-nom (p. ex. *la fille*) ou déterminant-adjectif (p. ex. *la petite*, Valois et al., 2009). Le syntagme nominal permet donc une évaluation très précoce du vocabulaire de l'enfant, en plus de fournir des données sur les opérations d'attribution du genre, d'accord de l'adjectif, et sur la maîtrise des éléments syntaxiques qui le composent. Une ébauche de la structure du syntagme nominal (adaptée de Ronat, 1977) est présentée dans la figure 1. On peut y voir que la structure du syntagme nominal permet une grande complexité de structures syntaxiques si plusieurs éléments (p. ex. des adjectifs et quantifieurs) y sont insérés.

Le genre grammatical (masculin/féminin) est une caractéristique importante du français. Il est utilisé pour établir la référence et maintenir la cohésion dans le discours et le texte. Les processus d'accord sont observés sur les pronoms clitiques (*il/elle, le/la*), certains participes passés (*mort/e*), les adjectifs (*vert/e*), ainsi que les pronoms toniques (*ceux-ci/celles-ci*, etc.). Les noms ont tous un genre en français[1] et tous les éléments du syntagme nominal doivent s'accorder entre eux ; ce processus est nommé *concordance* (Beauzée, 1767 ; Boloh & Ibernon 2010 ; Corbett, 2006). Le français a des adjectifs et des déterminants définis et indéfinis différents pour le masculin et le féminin au singulier (*le/la/un/une*). L'utilisation de déterminants se nomme *l'attribution* du genre (Boloh & Ibernon, 2010), car sa production est tributaire du genre qu'on attribue au nom. Leur forme est prévisible. L'*accord* du genre sur l'adjectif est irrégulier : la marque morphologique prend habituellement la forme de consonnes finales dites "flottantes" (p. ex. *vert-verte /vɛʁ-vɛʁt/*) qui apparaissent sur les formes féminines seulement (Paradis & El Fenne, 1995). Selon plusieurs auteurs, ces formes sont imprévisibles et ne peuvent être décrites ni *via* une règle d'élision (Paradis & El Fenne, 1995) ni *via* une insertion de consonne finale (Tranel, 2000). En effet, il n'y a pas de règle de formation du masculin à partir du féminin ni du féminin à partir du masculin (Fink, 1985 ; Herschensohn, 1993 ; Picard, 1996 ; Royle, 2011).

[1] Les noms épicènes comme *enfant* peuvent avoir les deux genres, et le déterminant est alors d'autant plus important pour en établir la référence.

Acquisition du concept du genre et de son accord

Valois et Royle (2009) ont observé dans un corpus longitudinal qu'une enfant francophone maîtrise très tôt l'attribution du genre du déterminant, et ce, dès l'âge de 18 mois (voir aussi Valois et al., 2009, pour des données de corpus transversaux d'enfants âgés de 20 à 36 mois). Les enfants peuvent produire spontanément des syntagmes nominaux avec une assignation du genre approprié (audible sur le déterminant) dès les premières combinaisons de mots (Valois & Royle, 2009 ; Valois et al., 2009). En contraste, certains enfants ayant un TDL peuvent encore omettre les déterminants en production spontanée à l'âge de 5 ans (Royle & Stine, 2013) ou montrent des difficultés à les produire sur des tâches induites telles que celle que nous présentons dans cet article (Royle & Reising, 2019).

Dans les corpus de production spontanée cités plus haut, on constate que l'inventaire spontané des adjectifs variables est limité. Il se pourrait que le contexte de l'échange n'ait pas généré d'occasion de les produire ou encore que les occasions qui se présentent ne sont pas des contextes où l'emploi de l'adjectif est obligatoire (pour éviter une rupture de communication par exemple) : les enfants éviteraient, ou ne verraient pas la nécessité, de produire des structures avec des adjectifs dans le syntagme nominal ou dans le syntagme verbal (p. ex., *elle est grande la maison*) et au final leur inventaire lexical spontané se limite à un très petit échantillon d'adjectifs variables : *grand, petit, gentil*, etc. Il est donc difficile d'évaluer la productivité et l'automatisation de processus d'accord. Néanmoins, on observe que les enfants à développement typique présentent des difficultés à produire les formes féminines variables (p. ex. *verte*) des adjectifs du français. Royle et Valois (2010) ont proposé qu'il n'existe pas de règle de flexion pour le féminin des adjectifs (voir aussi Royle, 2011, pour un argumentaire plus élaboré) et que ces items devaient être engrangés individuellement dans le lexique mental de l'individu, tout comme les verbes irréguliers. Soulignons qu'il existe des différences entre le positionnement syntaxique des adjectifs : certains sont prénominaux (p. ex. grandeur, *grand garçon*, ou âge, *vieille chaise*) tandis que d'autres sont postnominaux (p. ex. *robe verte*, etc.). Les enfants francophones maîtrisent précocement l'accord intra-nominal et le positionnement syntaxique des adjectifs en langage spontané dès l'âge de trois ans (Valois & Royle, 2009 ; Valois et al., 2009) tout en présentant des difficultés avec les formes féminines variables qu'ils mettent plus de temps à maîtriser (Jakubowicz et al., 1998 ; Jakubowicz & Roulet, 2007 ; Roulet-Amiot & Jakubowicz, 2006 ; Royle & Valois, 2010). Enfin, les jeunes avec un TDL font aussi des erreurs d'accord de l'adjectif, que ce soit en langage spontané ou en production induite. Ces faits ont été observés en français (Gopnik, 1990 ; Jakubowicz & Roulet, 2007 ; Roulet-Amiot & Jakubowicz, 2006 ; Royle & Reising, 2019 ; Royle & Stine, 2013), mais aussi en suédois (Leonard et al., 2001) et russe (Tribushinina & Dubinkina, 2012, voir Leonard, 2016 pour une revue des études sur le syntagme nominal).

Malgré leur acquisition précoce, les adjectifs (variables et invariables) ne sont pas produits en grande quantité dans les corpus de langage spontané d'enfants francophones (Valois & Royle, 2009 ; Valois et al., 2009) et ils sont moins bien représentés que les déterminants. Ceci est probablement dû à leur nature non-obligatoire et possiblement à leur redondance dans plusieurs contextes discursifs : un enfant ne mentionnera pas qu'un référent est vert dans un contexte où il n'y a qu'un seul référent possible (p. ex. *Donne-moi la voiture*). Pour cette raison, nous avons développé une tâche de production induite des adjectifs. De cette façon, nous pouvons contrôler les stimuli linguistiques produits par tous les enfants, mais aussi encourager les enfants à démontrer les capacités linguistiques qu'ils ne présenteraient pas en production spontanée.

L'objectif de cette étude est de décrire la performance d'enfants sur une tâche d'induction de la production de l'accord en genre sur les adjectifs et les déterminants. La tâche créée, dite « barrière », présente des casse-têtes avec des images d'objets et d'animaux de couleur et de taille variées.

Hypothèses : étant donné les études démontrant que l'accord sur les adjectifs variables, et surtout les adjectifs post-nominaux, reste difficile en production induite jusqu'à l'âge de 6 ans environ (Roulet-Amiot & Jakubowicz, 2006), nous nous attendons à observer un développement de la consolidation de l'accord de l'adjectif entre les âges de 3 à 6 ans, avec une maîtrise autour de l'âge de 6 ans. En contrepartie, l'attribution du genre avec le déterminant devrait être maîtrisée dès l'âge de 3 ans (Valois & Royle, 2009 ; Valois et al., 2009). La position de l'adjectif prénominal

devrait aussi se consolider plus tardivement que la postnominale (Roulet-Amiot & Jakubowicz, 2006). Enfin, la présence de deux adjectifs dans un syntagme (p. ex., *la grande grenouille verte*) devrait induire plus d'erreurs d'accord que la présence d'un seul adjectif (p. ex., *la grenouille verte*).

MÉTHODE

Participants

Deux études ont évalué au total 190 enfants (109 filles, 91 garçons). La majorité des enfants n'était pas régulièrement exposée à d'autres langues (que le français) et tous provenaient de la région de Montréal. Le tableau 1 présente des informations descriptives par groupe d'âge de 12 mois pour le premier groupe d'enfants de 3–6 ans (n = 155, Royle 2006–2008) et pour le second groupe de 6 à 9 ans (n = 35, Royle et al., 2009–2012). Les enfants de trois à six ans avaient entre 36 et 72 mois d'âge, avaient deux parents francophones, étaient exposés au français plus de 80% du temps et ne présentaient pas d'histoire familiale ou médicale pouvant nuire à leur développement. Ils ont été recrutés via des affiches déposées dans des lieux publics visant leurs parents (cliniques médicales et dentaires, pharmacies, magasins d'alimentation, centres sportifs, services de garde, etc.). Les enfants de plus de 72 mois avaient au moins un parent francophone, étaient scolarisés en français et utilisaient le français plus de 80% du temps. Ils ont été majoritairement recrutés dans une école privée proche du centre de recherche où se déroulait l'étude. Il y avait en moyenne 47–56 enfants par sous-groupe de 3 à 6 ans, et 12 enfants par groupe de 12 mois de 6 à 9 ans. Des informations démographiques sur les parents (niveau d'éducation) et les enfants (langues parlées, nombre de mois en garderie, rang dans la fratrie et taille de la fratrie) ont été recueillies via un questionnaire parental. Le protocole de recherche a été approuvé par le Comité d'éthique du Centre de recherche CHU Ste-Justine pour les deux études, et la seconde a également été approuvée par le comité d'éthique du Centre de recherche CRIUGM. La première étude visait le développement de l'accord chez les enfants en bas âge en langage spontané et en production induite et la seconde visait les bases neurocognitives de l'acquisition de l'accord à l'âge scolaire.

Déroulement de l'étude

Groupe 3–6 ans : Le parent ou tuteur était invité avec son enfant au centre de recherche du CHU Ste-Justine (Marie-Enfant) pour une rencontre d'environ une heure. Au cours de la rencontre, les enfants ont participé à une étude visant le développement de l'accord dans le syntagme nominal. Groupe 6–9 ans : Le parent ou tuteur était invité avec son enfant au CRIUGM pour deux rencontres d'environ une heure en l'espace de deux semaines. Au cours des deux rencontres, les enfants ont participé à une étude visant le développement des réponses neuronales (avec enregistrement de l'EEG - électroencéphalogramme) sur l'accord dans le syntagme nominal. Les tâches de casse-tête suivaient la seconde séance d'enregistrement.

Procédure

Tous les enfants ont été évalués avec des outils de dépistage orthophoniques et un dépistage auditif binaural. Un seuil auditif, dans les deux oreilles, supérieur à 20 dB à 500 Hz, ou à 15 dB à 1000, 2000 et 4000 Hz entraînait l'exclusion de l'enfant de l'étude[2].

Nous avons évalué les enfants de 3 à 6 ans sur leur vocabulaire réceptif avec une tâche normée pour le Canada (EVIP ; Dunn et al., 1993), la sous-tâche de mémoire non verbale (Leiter Memory Screen ; Roid & Miller, 1996), la production et la compréhension des adjectifs de couleur et de grandeur et l'accord en genre du déterminant défini et de l'adjectif de grandeur et de couleur avec une série de casse-têtes, la production du langage spontané au moyen d'un enregistrement audio-vidéo d'une période de jeu avec un ensemble standard de jouets (maison, train-cirque, pâte à modeler, autobus) (voir analyses de ce corpus dans Royle & Reising, 2019 ; Royle & Stine, 2013 ; Royle & Valois, 2010). Ces éléments ont été utilisés pour évaluer les structures d'intérêt et pour établir les scores de longueur moyenne d'énoncé (LME). Nous ne rapportons ici les résultats détaillés que pour la tâche de casse-têtes.

Chez les enfants de 6 à 9 ans, en plus des casse-têtes, les compétences linguistiques ont été vérifiées au moyen d'une évaluation du langage réceptif qui n'est pas normée pour le Canada (ÉLO ; Khomsi, 2001) et des tâches psycholinguistiques : une tâche de jugement de grammaticalité (Poulin et al., 2015), et trois tâches de définitude (Royle et al.,

[2] Suivant la recommandation de l'audiologiste, les enfants étaient retenus si une difficulté monaurale était observée. Seuls deux enfants de 3 à 6 ans n'ont pas été retenus suite à cette évaluation.

TABLEAU 1 : Caractéristiques des participants par sous-groupes d'âge, en moyennes (et écarts-types), sauf pour les colonnes « Nombre (*n*) » « Filles/ Garçons », et « Autre(s) langue(s) » où le nombre total est indiqué, et « Fratrie » ou l'étendue est présentée.

Groupe d'âge (intervalle)	n	Filles/ Garçons	Âge moyen (en mois)	Autre(s) langue(s)	Scolarité mère (en années)	Scolarité père (en années)	Scolarité parents (en années)	Garderie (en mois)	Rang	Fratrie	Fratrie *Min-Max*
3 (36–47)	56	33/23	41,88 (3,45)	10	15,93 [a] (2,89)	14,63 [b] (3,58)	15,24 [a] (2,77)	29,63 [a] (18,09)	1,52 (0,89)	2,09 (0,92)	1–5
4 (48–58)	50	31/19	53,06 (3,14)	16	15,25 (2,68)	14,47 [c] (3,09)	14,98 (2,50)	33,02 (14,18)	1,42 (0,73)	2,02 (0,82)	1–4
5 (59–72)	47	25/22	65,64 (4,29)	8	15,92 [a] (2,69)	15,51 [d] (3,07)	15,71 (2,66)	30,86 [a] (13,98)	1,36 (0,70)	2,13 (0,95)	1–5
6 (73–83)	12	5/7	77,52 (2,83)	5	17,42 (3,40)	16,91 (6,83)	17,17 (4,59)	16,00 (19,54)	2,00 (1,35)	2,67 (1,37)	1–6
7 (84–95)	13	8/5	91,39 (2,75)	3	18,00 (1,87)	15,42 (3,65)	16,81 (2,58)	27,15 (22,57)	1,45 (0,82)	2,08 (0,95)	1–4
8 (96–105)	12	7/5	99,75 (2,60)	5	18,27 (5,37)	17,82 (4,42)	17,58 (4,11)	21,75 (20,67)	2,00 (0,85)	2,58 (0,79)	1–4

Notes : [a.] 1 donnée manquante, [b.] 3 données manquantes, [c.] 5 données manquantes [d.] 4 données manquantes, n : nombre d'enfants dans le groupe ; Filles/Garçons : nombre de filles et de garçons dans le groupe ; Autre(s) langue(s) : Nombre d'enfants exposés à une (n = 39) ou deux (n = 7) autres langues ; Scolarité parents : scolarité moyenne des deux parents ; Garderie : nombre de mois en garderie ; Rang : rang moyen de l'enfant par rapport à sa fratrie ; Fratrie : taille moyenne de la fratrie ; Fratrie Min-Max : taille minimale et maximale de la fratrie.

2018), en plus d'une étude de neuro-imagerie avec potentiels évoqués sur la perception des erreurs de genre dans le syntagme nominal (Courteau et al., 2013).

Pendant les tâches de casse-tête (tâches avec barrière), la personne menant l'expérimentation installait un écran entre elle/lui et l'enfant pour éviter que l'enfant ne pointe du doigt la réponse. Les planches de casse-tête ont été présentées séquentiellement par ordre croissant de difficulté. Deux stimuli d'entraînement étaient présentés à l'enfant avant chaque casse-tête, avec des modèles cibles fournis oralement (par exemple, « Si tu voulais cette pièce, tu pourrais dire 'Donne-moi le cheval rose'. Je suis derrière l'écran et je ne peux pas voir le casse-tête, tu dois me dire avec des mots quelle pièce tu veux »). L'enfant était ensuite encouragé à demander la pièce qu'iel voulait mettre sur le casse-tête. S'il réussissait le premier casse-tête, l'adulte passait au second, et ainsi de suite. Ensuite, une tâche de compréhension a été effectuée en utilisant le même matériel et la même procédure, en inversant les rôles de l'adulte et de l'enfant.

Stimuli

Nous avons créé quatre casse-têtes contenant des images de noms dont la taille et la couleur variaient pour susciter des adjectifs ciblés. Tous les adjectifs et noms utilisés sont acquis précocement et sont de haute fréquence en français oral (New et al., 2001 ; Trudeau et al., 1999 ; voir Annexe A pour les détails sur les adjectifs et Annexe B pour des comparaisons avec les noms). Les images ont été imprimées sur le fond des planches du puzzle et sur des pièces de bois s'insérant dans la planche (voir exemple en Annexe C).

Le premier casse-tête (*Couleurs*) testait le vocabulaire des adjectifs de couleur : six adjectifs de couleur ont été utilisés, trois variables – *vert, blanc* et *brun*[3] – généralement produits au masculin par défaut (voir aussi les études psycholinguistiques de Boloh & Ibernon, 2010 et Violin-Wigent, 2006) et trois invariables – *jaune, bleu* et *rouge– rose* et *bleu*, étaient utilisés dans les exemples de familiarisation avec rétroaction corrective au besoin.

Le deuxième casse-tête (*Syntagmes Grandeurs* ou *SG*) présentait huit éléments : quatre noms (*canard, grenouille, maison* et *bateau*) dont la taille variait, ce qui obligeait à utiliser les adjectifs variables *petit* et *grand* (ou *gros*) pour les nommer. *Cheval* (*grand* et *petit*) était utilisé comme exemple de familiarisation. Les images étaient blanches avec un contour noir sur fond blanc.

Le troisième casse-tête (*Syntagmes Couleurs* ou *SC*) contenait les noms d'animaux (*canard* et *grenouille*) et les six couleurs du premier casse-tête, en plus des items de pratique *cheval rose* et *cheval bleu*. Les 12 combinaisons possibles ont été réparties entre deux versions, A et B. Étant donné que nous avions seulement trois adjectifs variables au total, ceci a créé un déséquilibre entre les versions A et B. Les enfants de 3 à 6 ans ont fait une des deux versions, tandis que les enfants de 6 à 9 ans ont fait les deux. Nous tiendrons compte de ce facteur dans les analyses.

Le quatrième casse-tête (*Syntagmes Grandeurs et Couleurs* ou *SGC*) présentait les noms *bateau* et *maison* avec des caractéristiques de couleur et de taille afin d'induire des structures avec deux adjectifs, un prénominal et un postnominal (par exemple, *le petit bateau rouge*) en plus d'images pour *grand cheval rose* et *petit cheval bleu*, items de familiarisation. Ces combinaisons résultent en 24 items répartis sur deux versions, A et B. Les enfants de 3 à 6 ans ont fait une des deux versions, tandis que les enfants de 6 à 9 ans ont fait les deux.

ANALYSES

Réponses cibles

Des analyses descriptives ont été faites sur les réponses cibles en production de structures complètes, c'est-à-dire l'adjectif seul (premier casse-tête seulement) ou le syntagme avec le déterminant défini, le nom et l'adjectif accordé. Dans le cas du premier casse-tête, on visait la production de la couleur cible, sans égard pour l'utilisation d'un déterminant. La dénomination d'adjectifs de façon isolée suscite généralement une réponse sous la forme masculine, mais une forme féminine était acceptée. Dans le cas des trois autres casse-têtes, on visait la production d'un déterminant défini (*le* ou *la*) et l'adjectif invariable ou variable dans la position syntaxique attendue (prénominale ou postnominale) dans sa forme féminine ou masculine, le cas échéant. Un point était donné pour la production complète de chaque item, avec tous les éléments obligatoires (déterminant et adjectif(s)). Si l'enfant se corrigeait

[3] "Marron" n'est généralement pas utilisé au Canada, sauf par des immigrants francophones européens.

et produisait une réponse cible, le point était donné. En revanche, toute erreur était notée pour les analyses qualitatives. Dans le premier casse-tête le score était sur 6, dans le second sur 8, dans le troisième sur 6 ou 12 (si les deux versions étaient faites) et dans le quatrième sur 12 ou 24. Dans le cas où les items avaient des formes masculines et féminines, les analyses incorporent le facteur genre (masculin, féminin). Les comparaisons ont aussi tenu compte du type d'adjectif (épicène, variable).

Analyses de productions non-cibles

Une analyse des productions non-cibles (grammaticales ou non) a aussi été faite. Étant donné le peu d'erreurs en compréhension, ces données n'ont pas été analysées plus avant. Les erreurs de production au cours de la première tâche ne concernaient que la production d'une autre couleur, par exemple *mauve* pour *brun*, ou l'absence de réponse.

Au cours des trois autres tâches, nous avons observé des erreurs d'attribution du genre sur le déterminant ou d'accord en genre sur l'adjectif (p. ex., *la grenouille *vert* ou *la petit bateau blanc*), des commissions (ou substitutions) d'adjectifs (p. ex., *maman* pour *grande* ou *grise* pour *brune*), l'omission de déterminants (p. ex., __ *canard rouge*) et la commission du déterminant (p. ex., **un grenouille rouge*). Les enfants pouvaient aussi produire des syntagmes fractionnés (p. ex., *le petit blanc, le bateau*) qui dans la plupart des cas étaient grammaticaux mais n'avaient pas la structure cible, et d'autres structures qui seront détaillées dans les sections suivantes. Dans les cas où il y avait plus d'une erreur dans la production, toutes les erreurs étaient comptabilisées pour l'analyse de réponses non-cibles, mais une seule était comptée pour la réussite.

Résultats

Les résultats aux tâches cognitives et linguistiques sont présentés dans le tableau 2, confirmant que les participants correspondaient au niveau attendu pour leur âge.

Pour chacune des tâches, les tableaux descriptifs suivants présentent la moyenne et l'écart type de chaque groupe, ainsi que deux autres scores pour interpréter les résultats. Dans un premier temps, les seuils de risque sont proposés en deçà desquels on peut considérer que l'enfant présente des difficultés. Ceux-ci ont été établis à un écart type sous la moyenne. Dans un deuxième temps, nous avons fixé un seuil de réussite de 80% à la tâche quel que soit l'âge des participants (voir Valois & Royle, 2009 pour l'acquisition de la première langue et Pallotti, 2007, pour l'acquisition de la

TABLEAU 2 : Résultats sur les tâches de dépistage des participants par sous-groupes d'âge, en moyennes (et écarts-types), sauf pour la colonne « Audio » où le nombre total est indiqué.

Groupe d'âge	QI mémoire	ÉVIP	LMÉw	LMÉm	Audio	ÉLO CI	ÉLO CG	Gramm
3 (36–47)	102,41[a] (15,10)	71,69 (28,53)	4,24 [b] (0,91)	5,60 [b] (1,19)	50 [a]	NA	NA	NA
4 (48–58)	109,94 [a] (17,34)	80,89 (21,72)	4,86 [c] (0,78)	6,46 [c] (1,02)	47	NA	NA	NA
5 (59–72)	110,19 (11,78)	90,97 (12,62)	5,36 [d] (0,68)	7,13 [d] (0,89)	47	NA	NA	NA
6 (73–80)	NA	NA	NA	NA	12	56,36 (23,67)	47,50 (23,60)	2,33 (1,72)
7 (84–95)	NA	NA	NA	NA	13	69,23 (27,22)	61,54 (16,51)	3,15 (1,28)
8 (96–105)	NA	NA	NA	NA	12	60,83 (19,75)	69,17 (21,41)	3,08 (1,51)

Notes : [a.] 2 données manquantes, [b.] 27 données manquantes [c.] 22 données manquantes [d.] 20 données manquantes. QI mémoire : score composite de la sous-tâche mémoire du Leiter (Roid & Miller, 1996) ; ÉVIP : rang centile de l'*Échelle de vocabulaire en images Peabody* (Dunn et al., 1993) ; LMÉw : longueur moyenne d'énoncés en mots ; LMÉm : longueur moyenne d'énoncés en morphèmes ; Audio : nombre d'enfants ayant réussi le dépistage audiologique bilatéral ; ÉLO CI : évaluation du langage oral, compréhension immédiate ; ÉLO CG : évaluation du langage oral, compréhension globale (Khomsi, 2001) ; Gramm : jugement de grammaticalité sur des erreurs d'accord du déterminant ou de l'adjectif.

langue seconde). La dernière colonne des tableaux rapporte la proportion des enfants de chaque âge ayant obtenu un score égal ou supérieur à 80%, ce que nous considérons comme représentant la proportion des enfants de chaque groupe ayant réussi la tâche.

Tâche 1 : Dénomination de couleurs

Réponses cibles sur la tâche Couleurs

Les données descriptives de la tâche Couleurs sont présentées dans le tableau 3. On observe que la majorité des enfants réussissent la dénomination de couleurs dès l'âge de 36 mois. La moyenne de réussite globale est de 5,52/6 (Médiane, 5 ; ET = 0,79). Dans les groupes de 4 ans et plus, seuls deux enfants n'ont pas au moins 5 réponses cibles. Dans le groupe le plus jeune, on observe plus d'erreurs, surtout avec l'item *brun* qui est toutefois plus difficile pour tous les enfants (66,84% en moyenne, ET = 47,07%). Dans le tableau 3 on peut voir que le seuil de risque, établi à un écart-type sous la moyenne, est de 5 pour les groupes plus âgés, et de 4 pour le groupe plus jeune (3–4 ans). De plus, on peut voir qu'une performance de 80% ou plus est atteinte par tous les enfants de 5 ans et plus.

Analyse d'erreurs sur la tâche Couleurs

Une analyse des réponses alternatives données (tableau 4) révèle que, globalement, les enfants produisent de façon égale une absence de réponse (3,19%) et d'autres adjectifs par exemple *gris* ou *mauve* pour *brun* (5,16%) ; cependant, cette répartition change avec l'âge. En effet, l'absence de réponse n'est réellement présente que chez les enfants les plus jeunes, tandis qu'à partir de l'âge de 4 ans, les enfants ont plutôt tendance à produire des commissions (par exemple, *chocolat*, *gris*, *noir* ou *mauve* pour *brun*).

Toutes les erreurs portent sur l'item *brun* chez les enfants plus âgés. Chez les plus jeunes, il y a plus de variabilité mais les commissions sont en grande majorité observées pour *brun*. Les omissions, dont 8 sont observées pour *blanc*, s'observent seulement chez les deux groupes les plus jeunes.

TABLEAU 3 : Moyennes (sur 6), écarts-types, seuils de risque et performance sur la tâche *Couleurs* selon les groupes d'âge.

Groupe	Moyenne (ET standard)	Seuil de risque	≥ 80 % (5/6)
3–3;11	5,14 (1,14)	4	87,5%
4–4;11	5,54 (0,65)	5	96%
5–5;11	5,74 (0,44)	5	100%
6–6;11	5,75 (0,45)	5	100%
7–7;11	5,78 (0,44)	5	100%
8–8;11	5,75 (0,45)	5	100%

TABLEAU 4 : Distribution des types d'erreurs, en nombre et en pourcentages de réponses, sur la tâche *Couleurs* selon le groupe d'âge (entre parenthèses, le pourcentage d'enfants du groupe qui produisent chaque type de réponse).

Groupe	*n*	Aucune réponse/Omission	Commission
3–3;11	56	29, 8,6% (27%)	19, 5,7% (34%)
4–4;11	50	8, 2,7% (12%)	15, 5,0% (28%)
5–5;11	47	1, 0,4% (2%)	11, 3,9% (23%)
6–6;11	12	0	4, 5,5% (33%)
7–7;11	13	1, 1,3% (8%)	2, 2,6% (15%)
8–8;11	12	0	3, 4,2% (25%)
Total	190	39, 3,4% (12%)	54, 4,7% (28%)

Tâche 2 : Syntagmes nominaux avec adjectifs de grandeur

Réponses cibles sur la tâche Syntagmes Grandeurs

Les données descriptives de la tâche *Syntagmes Grandeurs* sont présentées dans le tableau 5. On observe que certains enfants peuvent réussir cette tâche dès l'âge de 36 mois, mais qu'un nombre important d'entre eux (~40%) ne la réussissent pas. L'habileté à produire ces structures se consolide vers l'âge de 50 mois. On n'observe aucun effet de genre grammatical sur la compétence à produire ces structures. Dans le tableau 5, on peut voir que le seuil de risque, établi à un écart-type sous la moyenne, est de 8 pour les groupes plus âgés, de 6 pour les 6–7 ans, de 5 pour les 4–5 et 5–6 ans et de 3 pour le groupe plus jeune (3–4 ans). De plus, on peut voir qu'une performance de 80% minimum est atteinte par tous les enfants de 7 ans et plus. Par ailleurs, un enfant de 6–7 ans a fait plusieurs erreurs de genre sur cette tâche (p. ex., **le petite canard*). Cet enfant n'a pas d'histoire de troubles d'apprentissage et est exposé 100% du temps au français.

Analyse d'erreurs sur la tâche Syntagmes Grandeurs

Les types de réponses non-cibles répertoriées sur la tâche Syntagmes Grandeurs sont a) le genre de l'adjectif – p. ex., *la grenouille vert* –, b) l'omission de l'adjectif, c) un syntagme nominal fractionné, ce qui est grammatical – p. ex., *le canard*, *le gros* pour *le gros canard* – d) l'ordre de l'adjectif – p. ex., *le canard gros* –, e) le genre du déterminant p. ex., *le grosse grenouille* – et f) l'omission du déterminant[4]. Les productions non-cibles les plus fréquentes sont des erreurs de genre d'adjectif, les syntagmes fractionnés et l'omission du déterminant (tableau 6). Les erreurs de genre sur l'adjectif et le déterminant ne disparaissent qu'à 7 ans, mais la production de syntagmes fractionnés et les erreurs d'ordre d'adjectif avec un adjectif de grandeur (p. ex., *le canard gros*) ainsi que les erreurs d'omission ne devraient plus être observées à cet âge.

Tâche 3 : Syntagmes nominaux avec adjectifs de couleur

Réponses cibles sur la tâche Syntagmes Couleurs

Les données descriptives de la tâche *Syntagmes Couleurs* sont présentées dans le tableau 7 et la figure 2, ainsi que des tests statistiques de Student avec deux échantillons pour comparer les deux versions (A et B). Les résultats par groupe d'âges sont divisés selon la version (A ou B) administrée. Certains enfants peuvent réussir cette tâche dès l'âge de 37 mois, mais il y a une grande variabilité de taux de réussite et la version A (M = 4,57, ET = 1,49), qui contient plus de formes féminines variables, est moins bien réussie que la version B (M = 5,05, ET = 1,50, $t(2,26)$, 2,75, $p = 0,006$). En effet, le premier score de 100% de la version A ne s'observe qu'à 47 mois. De plus, on peut voir qu'une performance de 80% ou plus est atteinte par presque tous les enfants de 7 ans et plus.

Une comparaison des résultats de chaque groupe d'âge sur les tâches A et B a été faite afin de vérifier si les différences globales entre les tâches se répercutaient à travers les groupes d'âge. Pour les groupes de 3 à 5 ans, des tests F pour l'homogénéité de la variance ont été opérés puis des tests T unilatéraux pour deux exemples à variance égale, ou inégale, ont été opérés. Pour les groupes de 6 à 9 ans, des tests T unilatéraux par paires ont été opérés. Les résultats sont présentés dans le tableau 7. Des différences entre la version A et B, en faveur de la version B, apparaissent seulement dans les groupes 4 et 5 ans.

En toute cohérence avec les résultats globaux sur les tâches, l'habileté à produire ces structures diffère selon le genre grammatical (masculin *vs* féminin) et la variabilité de l'adjectif (épicène *vs* variable). Des ANOVAs ont été réalisées sur les résultats de chaque version de la tâche avec les facteurs fixes variabilité (variable *vs* épicène) et genre grammatical (masculin *vs* féminin). Les deux facteurs sont significatifs (variabilité : $F > 12$, $p < 0,001$ dans les deux cas ; genre grammatical : $F > 45$, $p < 0,001$ dans les deux cas) et interagissent dans les deux tâches (version A : $F(1,114)$ 56,32, $p < 0,0001$; version B : $F(1,112)$ 8,72, $p = 0,0038$). Les syntagmes sont significativement moins bien réussis sur les items féminins variables que sur tous les autres items, et ceci dans les deux tâches (voir figure 2).

De plus, l'adjectif de couleur lui-même peut être plus ou moins réussi, similairement à ce que nous avons vu dans la tâche 1. La forme féminine *brune*

[4] Les réponses erronées qui n'ont pas été répertoriées plus de 5 fois dans au moins un des groupes sont 1) la commission de l'adjectif – p. ex. *maman* pour *grand* (n = 5, ce type de réponse ne s'observe que dans le premier groupe), 2) la sur-régularisation, *grote* pour *grosse* (n = 1), 3) une structure relative (*le bateau qui est gros*), copule (*elle est grosse la maison*), génitive (*une grosse brune de maison*), prépositionnelle (*la maison en petit*), ou autre (n = 14).

TABLEAU 5 : Moyennes (sur 8), écarts-types, seuils de risque et performance selon les groupes d'âge sur la tâche *Syntagmes Grandeurs*.

Groupe	Moyenne (ET standard)	Seuil de risque	≥ 80 % (7/8)
3-3;11	5,40 (2,60)	3	59%
4-4;11	6,70 (1,89)	5	82%
5-5;11	7,04 (2,08)	5	93%
6-6;11	7,25 (1,60)	6	83%
7-7;11	8 (0)	8	100%
8-8;11	8 (0)	8	100%

TABLEAU 6 : Distribution des types d'erreurs, en nombre et en pourcentage de réponses, sur la tâche *Syntagmes Grandeurs* selon le groupe d'âge (entre parenthèses, le pourcentage d'enfants du groupe qui produisent chaque type de réponse).

Groupe	*n*	Genre adjectif	Omission adjectif	Syntagme fractionné	Ordre adjectif	Genre déterminant	Omission déterminant
3–3;11	56	47, 10,2% (38%)	13, 2,8% (7%)	49, 10,7% (38%)	10, 2,2% (9%)	9, 2% (9%)	40, 8,7% (18%)
4–4;11	50	21, 5,2% (22%)	1, 0,2% (2%)	13, 3,2% (20%)	5, 1,2% (8%)	15, 3,7% (20%)	9, 2,2% (14%)
5–5;11	47	8, 2,1% (21%)	3, 0,8% (6%)	7, 1,9% (9%)	4, 1,1% (4%)	10, 2,7% (17%)	18, 4,8% (11%)
6–6;11	12	7, 7,1% (25%)	0	1, 1% (8%)	0	7, 7,1% (25%)	0
7–7;11	13	0	0	0	0	0	0
8–8;11	12	0	0	0	0	0	0
Total	190	83, 18,0% (22 %)	17, 3,7% (1%)	70, 15,2% (19%)	19, 4,1% (5%)	43, 9,0 % (12 %)	67, 14,6% (1%)

TABLEAU 7 : Moyennes (sur 6), écarts-types et seuils de risque et performance selon les groupes d'âge sur la tâche *Syntagmes Couleurs*, selon la version A ou B.

	Version A				Version B					
Groupe	*n*	Moyenne (ET standard)	Seuil de risque	≥ 80 % (5/6)	*n*	Moyenne (ET standard)	Seuil de risque	≥ 80 % (5/6)	Test F	Test-T
3–4 ans	29	3,28 (1,80)	1	31%	26	3,92 (1,73)	2	46%	0,86	ns
4–5 ans	25	4,44 (0,90)	4	52%	27	5,04 (1,20)	4	81%	0,16	**0,027**
5–6 ans	24	4,67 (1,28)	4	75%	23	5,48 (1,10)	4	91%	0,48	**0,014**
6–7 ans	12	5,33 (0,62)	5	92%	12	5,42 (1,44)	4	83%	–	ns
7–8 ans	13	5,62 (0,62)	5	100%	13	5,85 (0,36)	6	100%	–	ns
8–9 ans	12	5,83 (0,37)	5	91%	12	5,92 (0,28)	6	100%	–	ns

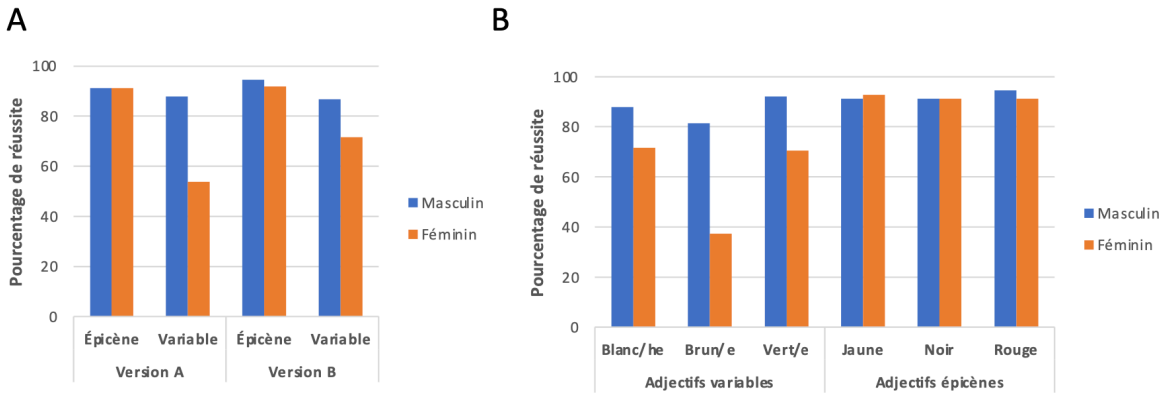


FIGURE 2 : Réussite en pourcentages de productions cibles du syntagme sur la tâche Syntagmes Couleurs selon le genre grammatical (féminin vs masculin) et le type d'adjectif (épicène vs variable) dans les versions A et B (A) et selon l'adjectif et le genre (B).

est la moins bien réussie (figure 2B). Étant donné que les formes brune et blanche sont toutes deux dans la version A de la tâche 3, il n'est donc pas étonnant qu'elle soit généralement moins bien réussie que la B qui ne contient que la forme féminine variable *verte*.

Analyse d'erreurs sur la tâche Syntagmes Couleurs

Dans le tableau 8, nous présentons le type de réponses non-cibles répertoriées sur les syntagmes avec adjectifs de couleur. Ceux-ci sont a) l'erreur de genre sur l'adjectif, b) l'omission de l'adjectif, c) la commission de l'adjectif d) l'ordre de l'adjectif – p. ex., *le canard gros* –, e) l'omission du déterminant, f) une erreur de définitude – p. ex., une grenouille blanche, g) un syntagme nominal fractionné – p. ex., *le canard, le brun* pour *le canard brun* – et h) la production d'une relative –p. ex., *la grenouille qui est blanche* (qui n'est observée que dans le groupe 3–4 ans[5]). Il est à noter que pour les enfants de 3–6 ans, le nombre d'erreurs est compilé sur 12 (sauf dans un cas) et dans le groupe 6–9 ans, sur 24.

Les productions non-cibles les plus fréquentes sont des erreurs de genre sur l'adjectif, les syntagmes fractionnés, l'omission ou la commission de l'adjectif et l'omission du déterminant. L'omission de l'adjectif disparait à 5 ans et les erreurs de définitude et la production de syntagmes fractionnés sont rares à cet âge. Les erreurs de genre sur l'adjectif diminuent après l'âge de 7 ans, mais cette erreur se retrouve particulièrement sur la forme *brune* et non pas *blanche* ou *verte*.

Tâche 4 : Syntagmes Couleurs et Grandeurs

Réponses cibles sur la tâche Syntagmes Couleurs et Grandeurs

Les données descriptives de la tâche *Syntagmes Couleurs et Grandeurs* sont présentées dans le tableau 9 ainsi que des tests statistiques de Student avec deux échantillons pour comparer les deux versions (A et B). Certains enfants peuvent réussir cette tâche dès l'âge de 43 mois, mais il y a une grande variabilité dans les taux de réussite. La version A (M = 8,48, ET = 3,58) semble généralement moins bien réussie que la version B (M = 9,38, ET = 3,53) mais cette tendance n'atteint pas le seuil de significativité dans une ANOVA généralisée incluant tous les groupes d'âge ($F(1,108)$ 1.86, p = 0,5, $t(1,218)$, 1,86, p = 0,06). Le premier score de 100% de la version A est observé à 44 mois et sur la forme B à 43 mois, mais les résultats sur les deux versions restent généralement faibles chez les enfants âgés de 3 et 4 ans. Enfin, on peut voir qu'une performance de 80% ou plus est atteinte par tous les enfants de 7 ans et plus sur la version B et par les enfants de 8 ans et plus sur la version A.

Plusieurs enfants du groupe âgés de 3 ans, 6 âgés de 36–38 mois, et 2 de plus de 42 mois, n'ont pas voulu faire cette tâche, ou n'ont réussi qu'à produire 2 ou 3 énoncés. Ces 8 enfants ont été retirés des résultats. Une comparaison des résultats de chaque groupe d'âge sur les tâches A et B a été

[5] Les réponses erronées qui n'ont pas été répertoriées plus de 5 fois dans au moins un des groupes sont 1) l'erreur de genre sur le déterminant (n = 5), 2), une structure copule (*elle est jaune la grenouille*) (n = 1) et 3) l'ordre de l'adjectif – p. ex, *le canard gros* – (n = 2).

opérée. On n'observe des différences significatives entre les versions A et B, en faveur de la version B, qu'aux âges de 5–6 ans.

Une vérification de l'effet du genre grammatical (masculin *vs* féminin) et de la variabilité (épicène *vs* variable) sur l'habileté à produire ces structures a été opérée. Étant donné que les adjectifs de grandeur étaient tous variables, le facteur variabilité n'a été pris en compte que pour les adjectifs de couleur[6]. Des ANOVAs ont révélé que les facteurs genre et variabilité ont un impact sur les résultats des deux versions (variabilité, $F(1, 108) > 14{,}57$, $p < 0{,}002$ dans les deux cas; genre grammatical : $F(1, 108) > 7{,}94$, $p < 0{,}005$ dans les deux cas) mais n'interagissent avec la variabilité que pour la version A (version A : $F(1{,}110)$ 31,68, $p < 0{,}0001$; version B : $F(1{,}108)$

[6] Voir la section des analyses d'erreurs pour plus de détails sur la réussite des différentes catégories d'adjectifs.

TABLEAU 8 : Distribution des types d'erreurs, en nombre et en pourcentage de réponses, sur la tâche *Syntagmes Couleurs* selon le groupe d'âge (entre parenthèses, le pourcentage d'enfants du groupe qui produisent chaque type de réponse).

Groupe	*n*	Genre adjectif	Omission adjectif	Commission adjectif	Omission déterminant	Définitude	Syntagme fractionné	Relative
3–3;11	55	50, 18,2% (64%)	19, 5,7% (24%)	14, 4,2% (22%)	20, 6,1% (12,7%)	3, 0,9% (4%)	38, 11,5% (33%)	8, 0,2% (5,5%)
4–4;11	50	38, 12,2% (60%)	7, 2,2% (8%)	9, 2,9% (16%)	5, 1,6% (2%)	11, 3,5% (4%)	10, 3,2% (14%)	0
5–5;11	47	26, 9,2% (45%)	0	4, 1,8% (11%)	12, 4,3% (8,5%)	3, 1,1% (2%)	1, 0,1% (2%)	0
6–6;11	12	6, 8,3% (42%)	0	3, 2,1% (25%)	9, 6,2% (16,7%)	0	0	0
7–7;11	13	5, 6,4% (31%)	0	3, 1,9% (15%)	0	0	0	0
8–8;11	12	2, 2,8% (17%)	0	0	0	0	0	0
Total	189	121, 11,6% (51%)	26, 2,4% (9%)	34, 3,1% (16%)	46, 4,2% (7,4%)	17, 1,6% (3%)	49, 4,5% (13%)	8, 0,7% (1,6%)

TABLEAU 9 : Moyennes (sur 12), écarts-types, seuils de risque et performance selon les groupes d'âge sur la tâche *Syntagmes Couleurs et Grandeurs*, selon la version A ou B.

	Version A				Version B					
Groupe	*n*	Moyenne (ET standard)	Seuil de risque	≥ 80 % (10/12)	*n*	Moyenne (ET standard)	Seuil de risque	≥ 80 % (10/12)	Test F	Test-T
3–3;11	25	6,08 (4,12)	2	28%	23	6,26 (3,80)	2	22%	0.36	ns
4–4;11	24	7,52 (2,39)	5	24%	26	8,12 (3,35)	5	38%	0,45	ns
5–5;11	24	8,38 (3,24)	5	42%	23	10,91 (1,93)	9	83%	0,01	**0,003**
6–6;11	12	9,17 (3,60)	9	67%	12	9,38 (4,02)	10	83%	–	0,10
7–7;11	13	11,23 (1,42)	10	85%	13	11,85 (0,53)	11	100%	–	ns
8–8;11	12	12 (0)	12	100%	12	12 (0)	12	100%	–	ns

Notes : 3 enfants de 36 mois, 2 de 37, 1 de 38, 1 de 42 et 1 de 46 mois n'ont pas participé ou n'ont pas pu répondre à suffisamment d'items à la tâche. 1 enfant de 4 ans ainsi que tous les enfants de 7–9 ans ont fait les deux versions de la tâche.

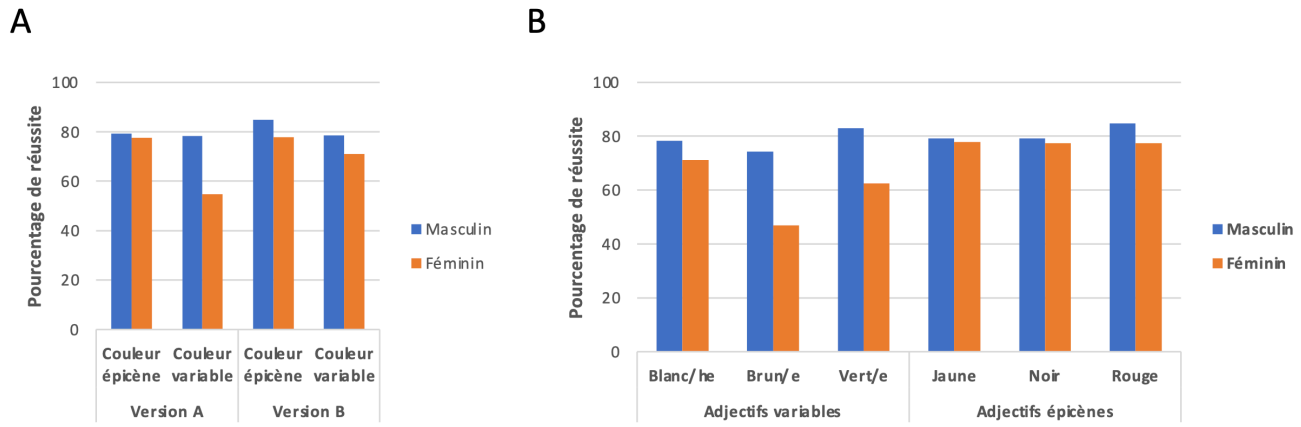


FIGURE 3 : Réussite en pourcentages de productions cibles du syntagme sur la tâche *Syntagmes Couleurs et Grandeurs* selon le genre grammatical (féminin *vs* masculin) et le type d'adjectif de couleur (épicène *vs* variable) dans les versions A et B (A), et selon l'adjectif et son genre (B).

0,019, p = 0,89). L'interaction reflète le fait que la forme *brune* se retrouve dans la version A : cette forme au féminin est la moins bien réussie (figure 3 B). Les syntagmes sont significativement moins bien réussis sur les items féminins variables que sur tous les autres items dans la tâche A. Dans la tâche B, ce sont généralement les formes féminines qui sont les moins bien réussies avec un effet additif de la variabilité sur le genre (voir figure 3A).

Analyse d'erreurs sur la tâche Syntagmes Couleurs et Grandeurs

Les types de réponses non-cibles répertoriées sur les syntagmes avec adjectifs de couleur et de grandeur sont les suivants : a) le genre d'adjectif de grandeur, b) le genre d'adjectif de couleur, c) l'omission de l'adjectif, d) la commission de l'adjectif, e) l'omission du déterminant, f) une erreur de définitude – p. ex., *une grenouille blanche*, g) l'ordre de l'adjectif – p. ex, *le canard gros brun* –, h) un syntagme nominal fractionné – p. ex., *le gros canard*, *le brun* pour *le gros canard brun* –, i) une relative –p. ex., *la grosse grenouille qui est blanche* et enfin j) une copule (p. ex., *le bateau noir, y'est petit*).[7] Les réponses de type « non-syntaxiques » sont présentées dans le tableau 10, et les « syntaxiques » dans le tableau 11.

Les productions non-cibles les plus fréquentes sont des syntagmes fractionnés, des erreurs d'ordre d'adjectifs (tableau 11), des erreurs de genre sur les adjectifs de couleur ou de grandeur, l'omission du déterminant et la commission de l'adjectif de couleur ou de grandeur (tableau 10). Les erreurs de genre sur l'adjectif restent présentes même après l'âge de 7 ans et cette erreur se retrouve presqu'exclusivement sur la forme *brune*. Le fractionnement du syntagme (qui est grammatical) est une stratégie très courante chez les jeunes enfants mais est rare ou disparait après 6 ans. Des erreurs d'ordre d'adjectif s'observent chez un sous-ensemble d'enfants jusqu'à 6 ans. Les relatives ne s'observent que chez 4 enfants, majoritairement du groupe 3 ans. L'omission et la commission d'adjectifs ou de déterminants ne devraient pas s'observer après 6 ans ; les erreurs de définitude et la production de relatives ou de copules après 4 ans sont rares et présentes chez une minorité d'enfants.

DISCUSSION

Accord du genre et assignation de genre

Dans des études longitudinales (Valois & Royle, 2009) et transversales (Valois et al., 2009), la maîtrise précoce de l'accord en genre est apparente chez les jeunes enfants francophones. Ceci signifie que ceux et celles-ci peuvent produire spontanément des syntagmes nominaux avec une assignation de genre approprié (visible sur le déterminant) dès les premières combinaisons de mots. Toutefois, l'emploi spontané limité d'adjectifs ne permet pas de statuer sur la maîtrise de l'accord de cette catégorie de mots. De plus, l'inventaire lexical spontané des enfants se limite à un très petit échantillon d'adjectifs variables : *grand*, *petit*, *gentil*,

[7] Les réponses erronées qui n'ont pas été répertoriées plus de 5 fois dans au moins un des groupes sont 1) la surrégularisation (p. ex., *noirte*, n = 3), 2), une structure conjointe (*la grenouille grosse et brune*) (n = 5), génitive (*la grande maison de brune*, n = 1) ou prépositionnelle (*le petit bateau comme le soleil*, n = 7)

etc. Il est donc difficile d'évaluer la productivité et l'automatisation de processus d'accord en dehors des tâches qui les ciblent spécifiquement. Les casse-têtes que nous avons créés ont été conçus à cet effet et nos données permettent d'identifier les âges charnières auxquels les enfants devraient maîtriser ces structures, c'est-à-dire qu'on en connaît l'âge attendu de maîtrise (i.e., 80% ou plus).

Les résultats actuels suggèrent que les adjectifs de grandeur sont mieux maîtrisés que les couleurs. En effet, dans les syntagmes à un seul adjectif, aucune erreur d'accord n'est notée à partir de 7 ans pour les items ciblant les adjectifs de grandeur. A contrario, certains items ciblant les adjectifs de couleur provoquent des erreurs, notamment « brun » et « blanc ». La maîtrise plus rapide des adjectifs de grandeur dans nos tâches est possiblement liée aux contraintes sémantiques sur les catégories d'adjectifs qui peuvent être dans une position prénominale (voir Valois et al., 2009 pour des détails). Le fait que la tâche de grandeur n'ait présenté que deux adjectifs, qui sont tous deux rapportés comme étant acquis précocement et même présents dans les corpus spontanés, peut aussi expliquer en partie cette observation.

TABLEAU 10 : Distribution des types d'erreurs non syntaxiques, en nombre et en pourcentage de réponses, sur la tâche *Syntagmes Couleurs et Grandeurs* selon le groupe d'âge (entre parenthèses, le pourcentage d'enfants du groupe qui produisent chaque type de réponse).

Groupe	*n*	Genre adjectif grandeur [a]	Genre adjectif couleur [a]	Omission adjectif grandeur ou couleur	Commission adjectif grandeur ou couleur	Omission déterminant	Définitude
3–3;11	48	52, 8,7% (44%)	72, 12,5% (65%)	23, 4% (17%)	46, 8,0% (40%)	31, 5,4% (15%)	7, 1,2% (4%)
4–4;11	50	22, 3,7% (16%)	66, 11,0% (56%)	5, 0,3% (6%)	22, 3,7% (24%)	14, 2,3% (6%)	12, 2,0% (4%)
5–5;11	47	6, 1,1% (11%)	29, 5,1% (29%)	2, 0,4% (2%)	8, 1,4% (15%)	21, 3,7% (13%)	0
6–6;11	12	4, 1,4% (25%)	5, 1,7% (25%)	2, 0,7% (8%)	8, 2,8% (25%)	45, 15,5% (16,7%)	0
7–7;11	13	5, 1,6% (15%)	6, 1,9% (15%)	0	0	0	0
8–8;11	12	2, 0,7% (17%)	0	0	2, 0,7% (8,3%)	0	0
Total	182	91, 3,5% (23%)	178, 6,8% (43%)	32, 1,2% (10%)	86, 1,0% (23%)	111, 4,2% (10%)	19, 0,1% (10%)

Notes : [a] Ce type d'erreurs ne s'observe que sur les adjectifs variables

TABLEAU 11 : Distribution des types d'erreurs syntaxiques, en nombre et en pourcentage de réponses, sur la tâche *Syntagmes Couleurs et Grandeurs* selon le groupe d'âge (entre parenthèses, le pourcentage d'enfants du groupe qui produisent chaque type de réponse).

Groupe	*n*	Syntagme fractionné	Ordre de l'adjectif	Relative	Copule
3–3;11	48	145, 24,3% (67%)	21, 3,5% (19%)	23, 3,9% (6,3%)	7, 1,2% (10,4%)
4–4;11	50	100, 16,2% (70%)	17, 2,8% (14%)	0	4, 0,6% (6%)
5–5;11	47	35, 6,2% (38%)	19, 3,3% (15%)	3, 0,5% (2,1%)	2, 0,4% (2,1%)
6–6;11	12	14, 4,7% (33%)	3, 1% (17%)	0	1, 0,3% (8,3%)
7–7;11	13	0	0	0	0
8–8;11	12	0	0	0	0
Total	182	292, 11,2% (49%)	60, 2,3% (14%)	26, 0,2% (2%)	14, 0,5% (6%)

Dans les syntagmes à deux adjectifs, des erreurs persistent sur l'accord en genre des adjectifs de couleur jusqu'à 8 ans et des adjectifs de grandeur jusqu'à 9 ans. La complexification de la structure pourrait a priori rendre l'accord en genre plus fragile, même pour des adjectifs qui semblaient bien maîtrisés lorsqu'employés dans un syntagme nominal plus simple. De plus, d'autres études démontrent que l'accord de verbes (Courteau et al., 2024 ; Franck et al., 2004) ne s'automatise pas avant l'âge de 8 ans, ce qui pourrait signifier un délai typique dans l'automatisation de processus linguistiques de vérification grammaticale de l'accord. Néanmoins, en comparant les taux d'erreurs pour chaque type d'adjectif dans la tâche la plus complexe (tableau 11), on constate que le taux d'erreur sur les couleurs est généralement plus élevé que pour les grandeurs, confirmant les résultats exposés précédemment.

Il faut noter que l'erreur de genre sur un adjectif de couleur sur ces tâches était attendue pour l'enfant qui apprend le français en L2, même si la L1 contient des traits de genre (par exemple, voir des données de l'espagnol et du français, Royle et al., 2015). À la différence des enfants avec un TDL, les enfants hispanophones apprenant le français langue seconde produisent proportionnellement plus d'erreurs de genre sur les adjectifs de grandeur, mais aussi une quantité importante d'erreurs d'inversion syntaxique de l'adjectif, en cohérence avec la grammaire de l'espagnol (p. ex., *le bateau grand*). La majorité des adjectifs en français sont postnominaux, comme en espagnol et de ce fait, on s'attendrait à trouver des erreurs de syntaxe du syntagme nominal qui suivrait cette tendance, mais ce n'est pas ce qui est observé en spontané en français (Valois & Royle, 2009). Il se pourrait que cette absence d'erreurs soit tributaire de l'importance que les adjectifs prénominaux ont pour la référence : ils identifient des traits sémantiques comme la grandeur et l'âge, et ce sont des adjectifs qui permettent notamment l'effacement du nom (p. ex., *je veux la grande*, Valois & Royle, 2009). Leur validité de signal (*cue validity*)[8] pourrait en promouvoir l'apprentissage rapide. Malgré leur importance pour la référence, on observe un nombre relativement fréquent d'erreurs d'ordre de l'adjectif postnominal avant l'âge de 5 ans sur la tâche 2 (*Grandeurs*) et jusqu'à l'âge de 6 ans sur la tâche la plus complexe (*Couleurs et Grandeurs*). Ce résultat démontre qu'à la différence des corpus de langage spontané, la tâche d'induction des adjectifs est sensible à cette différence de traitement des adjectifs pré- et postnominaux.

Enfin, il est intéressant de comparer les données d'enfants à développement typique (DT) à celles d'enfants présentant un TDL et de vérifier si les normes proposées sont utiles pour l'identification du TDL. Des données de 8 enfants ayant un TDL et âgés entre 5 et 5;11 ans (Royle & Stine, 2013 ; Royle & Reising, 2019), ainsi qu'un enfant de 6;06 ans qui n'avait pas été inclus dans les études précédentes étant donné son âge, sont présentées dans le tableau 12. On observe que les enfants avec un TDL réussissent habituellement la dénomination de couleurs et que peu d'enfants échouent sur la tâche 2 (*Grandeurs*) ou 3 (*Couleurs*) ;

[8] C'est-à-dire « la valeur informative d'un dispositif linguistique donné en tant qu'indice d'un sens ou d'une intention sous-jacente » (traduction libre de Bates et MacWhinney 1989, p. 37).

TABLEAU 12 : Résultats d'enfants francophones avec TDL sur les quatre tâches, basés sur les seuils établis.

Enfant	Âge (en mois)	Groupe d'âge	T1	T2	T3 A	T3 B	T4 A	T4 B
TDL1	80	6–6;11	Réussite	Réussite	–	**Échec**	–	**Échec**
TDL2	65	5–5;11	Réussite	**Échec**	Réussite	–	Réussite	–
TDL3	70	5–5;11	Réussite	Réussite	Réussite	Réussite	Réussite	**Échec**
TDL4	60	5–5;11	**Échec**	**Échec**	**Échec**	–	**Échec**	–
TDL5	67	5–5;11	Réussite	Réussite	Réussite	Réussite	Réussite	**Échec**
TDL6	67	5–5;11	Réussite	**Échec**	Réussite	Réussite	Réussite	**Échec**
TDL7	69	5–5;11	Réussite	Réussite	**Échec**	**Échec**	**Échec**	**Échec**
TDL8	69	5–5;11	Réussite	Réussite	Réussite	Réussite	Réussite	Réussite
TDL9	71	5–5;11	Réussite	Réussite	Réussite	Réussite	Réussite	Réussite

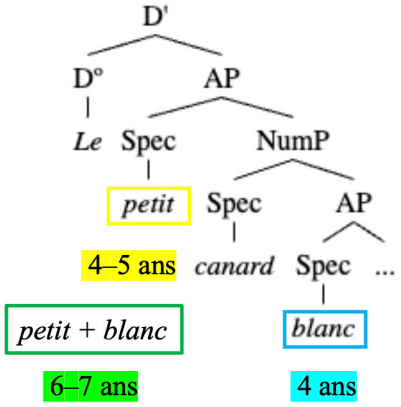


FIGURE 4 : Évolution des compétences en syntaxe du syntagme nominal chez l'enfant francophone.

cependant la majorité (7 sur 9) échouent à la tâche 4 B (*Couleurs et Grandeurs*). Ces enfants ont tendance à produire plus d'erreurs de genre sur les adjectifs variables que leurs pairs de même âge. Un autre schème observé chez un enfant avec un TDL est la surutilisation d'adjectifs féminins au lieu de masculins. Ce comportement n'a *jamais* été observé chez les 190 témoins de l'étude. Certains enfants TDL produisent des erreurs de genre sur les déterminants ou les omettent, comportements rares chez les témoins à cet âge. Ces données semblent indiquer que les enfants avec TDL, en dépit de leur bonne connaissance du vocabulaire utilisé, présentent des difficultés disproportionnées avec les processus d'accord du français, surtout de l'accord adjectif-nom mais aussi parfois de l'attribution du genre sur le déterminant, comportement qui n'est pas observé en langage spontané (Royle & Reising, 2019).

La syntaxe du syntagme nominal

Cette étude nous a permis d'évaluer certains aspects du développement de la maîtrise de la syntaxe du syntagme nominal chez l'enfant francophone. Des études de corpus ont montré que la production de syntagmes nominaux avec des adjectifs est rare dans le discours spontané des enfants (Valois & Royle, 2009 ; Valois et al., 2009) et que les structures complexes avec des adjectifs ne représentent que 8,5 à 12 % de ce discours (Royle & Stine, 2013). En effet, dans les corpus, les structures avec des adjectifs dans le syntagme nominal ou dans le syntagme verbal (p. ex., *elle est grande la maison*) sont peu variées et peu fréquentes. Ainsi, la tâche proposée ici constitue un outil grâce auquel les cliniciens pourront observer la production de différents degrés de complexité du syntagme nominal et en évaluer la maîtrise en présentant des occasions de production contrôlées avec des noms et adjectifs connus des enfants.

Nous représentons schématiquement dans la figure 4 nos résultats avec une version simplifiée de l'arbre syntaxique présenté dans la figure 1. La première structure maîtrisée par les enfants est l'adjectif postnominal (*couleur*) dès l'âge de 4 ans : il n'y a pas d'erreurs d'ordre d'adjectifs et peu de fractionnement. Les structures avec des adjectifs de grandeur (prénominaux) sont acquises juste après, soit entre 4 et 5 ans. Les erreurs d'ordre d'adjectif sont rares après 4 ans et les syntagmes fractionnés le sont après 5 ans. Enfin, les structures plus complexes impliquant les deux types d'adjectifs sont maîtrisées plus tardivement, entre 6 et 7 ans. On observe que les erreurs de fractionnement pour ces structures plus complexes sont encore présentes à 6 ans ainsi que des erreurs d'ordre d'adjectifs. Ces dernières se retrouvent sur les adjectifs prénominaux qui sont majoritairement produits à la fin du syntagme, par ex., *le bateau noir grand*. Ceci appuie l'interprétation selon laquelle les enfants n'ont pas encore bien maîtrisé la position syntaxique de ces adjectifs, surtout dans des structures demandant plus d'un type d'adjectif. Néanmoins, la majorité des enfants commettant cette erreur (n =19) la font entre une et trois fois au cours de la tâche et seuls six enfants font entre 5 et 8 erreurs de ce type.

CONCLUSION

Cette étude avait pour but de décrire la production de syntagmes nominaux à un ou deux adjectifs dans une série de tâches graduées, afin de documenter l'acquisition de structures variées du syntagme

nominal et de l'accord en genre de l'adjectif. Les résultats rapportés permettent de situer la performance d'un enfant par rapport à ses pairs de multiples façons. Tout d'abord, il est possible de statuer sur l'atteinte du seuil critique établi pour chaque tâche et chaque groupe d'âge. Un enfant sous le seuil critique se trouverait dans les 10% les plus faibles de son groupe d'âge. Ensuite, il est possible de statuer sur la réussite d'une tâche avec un seuil fixe de 80%, permettant ainsi d'établir la probabilité pour un enfant de réussir la tâche en fonction de son âge. Finalement, une analyse des erreurs peut être réalisée afin de voir si, au-delà de leur nombre, les erreurs produites par l'enfant sont typiques pour son âge ou pas. En alliant des points de repère quantitatifs et qualitatifs, l'analyse des productions de l'enfant s'en trouvera enrichie et pourra alimenter le raisonnement des orthophonistes menant à la conclusion clinique. D'autres travaux sont évidemment requis afin de renforcer nos connaissances sur la maîtrise de ces structures, notamment auprès d'enfants allophones ou présentant des troubles ou des retards d'apprentissage du langage.

DECLARATION D'INTÉRETS

Les autrices ont déclaré n'avoir aucun lien d'intérêt en relation avec cet article.

RÉFÉRENCES

Bates, E., & MacWhinney, B. (1989). Functionalism and the competition model. Dans B. MacWhinney et E. Bates (dir.), *The Crosslinguistic Study of Sentence Processing* (p. 3–73). Cambridge University Press.

Beauzée, N. (1767). Grammaire générale ou Exposition raisonnée des éléments nécessaires du langage : pour servir de fondement à l'étude de toutes les langues (livre III, t. II). Paris, Barbou.

Boloh, Y., & Ibernon, L. (2010). Gender attribution and gender agreement in 4- to 10-year-old French children. *Cognitive Development, 25*(1), 1–25. https://doi.org/10.1016/j.cogdev.2009.09.011

Bouchard, M. E., Fitzpatrick, E. M., & Olds, J. (2009). Analyse psychométrique d'outils d'évaluation utilisés auprès des enfants francophones. *Revue canadienne d'orthophonie et d'audiologie, 33*(3), 129-139.

Corbett, G. G (2006). Gender, grammatical. Dans K. Brown (dir.), *Encyclopedia of Language & Linguistics* (p. 749-756). Elsevier.

Courteau, É., Royle, P., Gascon, A., Marquis, A., Drury, J. E., & Steinhauer, K. (2013). Gender concord and semantic processing in French children: An auditory ERP study. Dans S. Baiz, N. Goldman, & R. Hawkes (dir.), *BUCLD 37: Proceedings of the 37th annual Boston University Conference on Language Development* (vol. 1, p. 87-99). Cascadilla Press.

Courteau, É., Royle, P., & Steinhauer, K. (2024). Number agreement processing in adolescents with and without developmental language disorder (DLD): Evidence from event-related brain potentials. *Scientific Reports, 13*, 22836. https://doi.org/10.1038/s41598-023-49121-1

Dunn, L. M., Thériault-Whalen, C., & Dunn, L. M. (1993). *EVIP : Échelle de vocabulaire en images Peabody*. Psycan.

Filiatrault-Veilleux, P., Desmarais, C., Bouchard, C., Trudeau, N., & Leblond, J. (2016). Conception et qualités psychométriques d'un outil d'évaluation de la compréhension d'inférences en contexte de récit chez des enfants âgés de 3 à 6 ans. *Revue canadienne d'orthophonie et d'audiologie, 40*(2), 149–163.

Fink, R. (1985). French adjective morphophonemic patterns: Their generalization and representation. *Linguistics, 23*(4), 567–596. https://doi.org/10.1515/ling.1985.23.4.567

Franck, J., Cronel-Ohayon, S., Chillier, L., Frauenfelder, U. H., Hamann, C., Rizzi, L., & Zesiger, P. (2004). Normal and pathological development of subject-verb agreement in speech production: A study on French children. *Journal of Neurolinguistics, 17*(2–3), 147–180. https://doi.org/10.1016/S0911-6044(03)00057-5

Gopnik, M. (1990). Feature blindness: A case study. *Language Acquisition, 1*(2), 139–164. https://doi.org/10.1207/s15327817la0102_1

Grüter, T. (2005). Comprehension and production of French object clitics by child second language learners and children with specific language impairment. *Applied Psycholinguistics, 26*(3), 363–391. https://doi.org/10.1017/S0142716405050216

Herschensohn, J. (1993). Applying linguistics to teach morphology: Verb and adjective inflection in French. *International Review of Applied Linguistics in Language Teaching, 31*(2), 97–112. https://doi.org/10.1515/iral.1993.31.2.97

Jakubowicz, C., Nash, L., Rigaut, C., & Gérard, C.-L. (1998). Determiners and clitic pronouns in French-speaking children with SLI. *Language Acquisition, 7*(2–4), 113–160. https://www.jstor.org/stable/20000283

Jakubowicz, C., & Nash, L. (2001). Functional categories and syntactic operations in (ab)normal language acquisition. *Brain and Language, 77*(3), 321–339. https://doi.org/10.1006/brln.2000.2405

Jakubowicz, C. (2003). Computational complexity and the acquisition of functional categories by French-speaking children with SLI. *Linguistics, 41*(2), 171–211. https://doi.org/10.1515/ling.2003.007

Jakubowicz, C., & Roulet, L. (2007). Narrow syntax or interface deficit? Gender agreement in French SLI. Dans J. Liceras, H. Zobl et H. Goodluck (dir.), *The Role of Formal Features in Second Language Acquistion* (p. 184–225). Lawrence Erlbaum.

Khomsi, A. (2001). *ÉLO : Evaluation du langage oral*. ECPA.

Leonard, L. B. (2016). Noun-related morphosyntactic difficulties in specific language impairment across languages. *First Language, 36*(1), 3–29. https://doi.org/10.1177/0142723715626070

Leonard, L. B., Salameh, E.-K., & Hansson, K. (2001). Noun phrase morphology in Swedish-speaking children with specific language impairment. *Applied Psycholinguistics, 22*(4), 619–639. https://doi.org/10.1017/S0142716401004076

New, B., Pallier, C., Ferrand, L., & Matos, R. (2001). Une base de données lexicales du français contemporain sur internet : LEXIQUE. *L'Année Psychologique, 101*(3-4), 447–462. http://doi.org/10.3406/psy.2001.1341

OOAQ (2014). *Mémoire de l'Ordre des orthophonistes et audiologistes du Québec présenté à la Commission de la santé et des services sociaux relativement au projet de loi n°10. Loi modifiant l'organisation et la gouvernance du réseau de la santé et des services sociaux notamment par l'abolition des agences régionales.* https://www.assnat.qc.ca

Pallotti, G. (2007). An operational definition of the emergence criterion. *Applied Linguistics, 28*(3), 361–382. https://doi.org/10.1093/applin/amm018

Paradis, C., & El Fenne, F. (1995). French verbal inflection revisited: Constraints, repairs and floating consonants. *Lingua, 95*(1-3), 169–204. https://doi.org/10.1016/0024-3841(95)90105-1

Paul, M. J., & Thordardottir, E. T. (2019). Investigation of the psychometric properties of the "Milestones en français du Québec" (MilBec), a new language screener for French-speaking children between 12 and 71 months. *Canadian Journal of Speech-Language Pathology and Audiology, 43*(2), 133-152. https://www.cjslpa.ca/files/2019_CJSLPA_Vol_43/No_2/CJSLPA_Vol_43_No_2_2019_MS_1103.pdf

Pesco, D., & O'Neill, D. (2016). Assessing early language use by French-speaking Canadian children: Introducing the LUI-French. *Canadian Journal of Speech-Language Pathology and Audiology, 40*(3), 198-217. http://cjslpa.ca/files/2016_CJSLPA_Vol_40/No_03/CJSLPA_2016_Vol_40_No_3_Pesco_O_Neill_198-217.pdf

Picard, M. (1996). The empty onset principle: A problematic phonological constraint. *Revue québecoise de linguistique théorique et appliquée*, 13, 195–201.

Poulin, M.-J., Marquis, A., & Royle, P. (2015). Étude de faisabilité portant sur l'évaluation de la production et de la compréhension du langage oral en français. Dans M. Pomerleau & E.-M. Gendron-Pontbriand (dir.), *ScriptUM : revue de l'édition 2014 du colloque multidisciplinaire sur le langage VocUM* (n°1, p. 54–68). https://scriptum.vocum.ca/index.php/scriptum/article/view/28

Roid, G. H., & Miller, L. J. (1996) *The Leiter International Performance Scale-Revised.* Stoelting Co.

Ronat, M. (1977). Une contrainte sur l'effacement du nom. Dans M. Ronat (dir.) *Langue : théorie générative étendue* (p. 153-169). Herman.

Rose, Y., & Royle, P. (1999). Uninflected structure in familial language impairment: Evidence from French. *Folia Phoniatrica et Logopaedica, 51*(1-2), 70-90. https://doi.org/10.1159/000021482

Roulet-Amiot, L., & Jakubowicz, C. (2006). Production and perception of gender agreement in French SLI. *Advances in Speech-Language Pathology, 8*(4), 335-346. https://doi.org/10.1080/14417040601009420

Royle, P. (2006–2008). *Outils morphosyntaxiques pour le dépistage de la dysphasie chez les jeunes francophones.* Subvention FQRSC Établissement de nouveaux professeurs-chercheurs (2006-NP-104790).

Royle, P. (2007). Variable effects of morphology and frequency on inflection patterns of French preschoolers. *The Mental Lexicon Journal, 2*(1), 103-125. https://doi.org/10.1075/ml.2.1.07roy

Royle, P., & Thordardottir, E. T. (2008). Elicitation of the passé composé in French pre-schoolers with and without SLI. *Applied Psycholinguistics, 29*(3), 341-365. https://doi.org/10.1017/S0142716408080168

Royle, P., Valois, D., Stemmer, B., Steinhauer, K., Drury, J. E., & Jutras, B. (2009–2012). *Tracking the acquisition of syntax and semantics in French noun phrases: A theoretical, behavioral and electrophysiological approach.* Subvention CRSH (410-2009-0256).

Royle, P., & Valois, D. (2010). Acquisition of adjectives in Quebec French as revealed by elicitation data. *Journal of French Language Studies, 20*(3), 313–338. https://doi.org/10.1017/S0959269510000013

Royle, P., Toupin, J., Bourguignon, N., Trudeau, N., & Valois, D. (2010). L'accord de l'adjectif chez les enfants francophones avec un trouble spécifique du langage. *SPECTRUM, 2*, 1–18. https://eoa.umontreal.ca/wp-content/uploads/sites/32/spectrumVol2RoyleTSL.pdf

Royle, P. (2011). On the existence of C/Ø alternations in French adjectives: Theoretical and empirical questions. *Proceedings of the 17th ICPhS. 4 p*, Hong Kong, 1730-1733. https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2011/index.htm

Royle, P., & Stine, I. (2013). The French noun phrase in preschool children with SLI: morphosyntactic and error analyses. *Journal of Child Language, 40*(5), 945–970. https://doi.org/10.1017/S0305000912000414

Royle, P., Bergeron, E., & Marquis, A. (2015). L'acquisition du genre quand il est déjà acquis : l'expérience d'enfants hispanophones apprenant le français. *Revue canadienne d'orthophonie et d'audiologie, 39*(3), 238–259. https://hdl.handle.net/1866/19913

Royle, P., Fromont, L. A., & Drury, J. E. (2018). Definiteness and maximality in French language acquisition, more adult-like than you would expect. *Frontiers in Communication, 3*, 27. https://doi.org/10.3389/fcomm.2018.00027

Royle, P., & Reising, L. (2019). Elicited and spontaneous determiner phrase production in French speaking children with developmental language disorder. *Canadian Journal of Speech Language Pathology and Audiology, 43*(3), 167–187. https://www.cjslpa.ca/detail.php?ID=1249etlang=en

Thordardottir, E. T. (2005). Early lexical and syntactic development in Quebec French and English: Implications for cross-linguistic and bilingual assessment. *International Journal of Language & Communication Disorders, 40*(3), 243–278. https://doi.org/10.1080/13682820410001729655

Thordardottir E. T., & Namazi, M. (2007). Specific language impairment in French-speaking children: Beyond grammatical morphology. *Journal of Speech, Language, and Hearing Research, 50*(3), 698–715. https://doi.org/10.1044/1092-4388(2007/049)

Thordardottir, E., Keheyia, E., Lessard, N., Sutton, A., & Trudeau, N. (2010). Typical performance on tests of language knowledge and language processing of French-speaking 5-year-olds. *Canadian Journal of Speech-Language Pathology and Audiology, 34*(1), 5–16. https://www.cjslpa.ca/detail.php?ID=996&lang=en

Tranel, B. (2000). Aspects de la phonologie du Français et la théorie de l'optimalité. *Langue française*, 126, 39–72. https://doi.org/10.3406/lfr.2000.989

Tribushinina, E., & Dubinkina, E. (2012). Adjective production by Russian-speaking children with specific language impairment. *Clinical Linguistics & Phonetics, 26*(6), 554–571. https://doi.org/10.3109/02699206.2012.666779

Trudeau, N., Frank, I., & Poulin-Dubois, D. (1999). Une adaptation en français du MacArthur Communicative Development Inventory. *Revue canadienne d'orthophonie et d'audiologie, 23*(2), 61–73. https://eoa.umontreal.ca/agora-des-professionnels/ressources/inventaires-macarthur-bates-imbdc/

Trudeau, N., & Sutton, A. (2011). Expressive vocabulary and early grammar of 16- to 30-month-old children acquiring Quebec French. *First Language, 31*(4), 480–507. https://doi.org/10.1177/0142723711410828

Valois, D., & Royle, P. (2009). Partitivity, atomization, and N-Drop: A longitudinal study of French child language. *Language Acquisition, 16*(2), 82–105. https://doi.org/10.1080/10489220902773079

Valois, D., Royle, P., Sutton, A., & Bourdua-Roy, E. (2009). L'ellipse du nom en français : le rôle des données de l'acquisition pour la théorie linguistique. *Revue canadienne de linguistique, 54*(2), 339–366. https://doi.org/10.1353/cjl.0.0050

Violin-Wigent, A. (2006). Gender assignment to nouns codeswitched into French: Observations and explanations. *International Journal of Bilingualism, 10*(3), 253-276. https://doi.org/10.1177/13670069060100030101

Wiig, E. H., Semel, E., & Secord, W. A. (2019). *Évaluation clinique des notions langagières fondamentales - Cinquième édition: Version pour francophones du Canada (CELF-5-CDN-F)*. Pearson. https://www.pearsonclinical.fr/celf-5-batterie-d-evaluation-des-fonctions-langagieres-et-de-communication

ANNEXE A : Fréquence des adjectifs de couleur selon la base de données Frantext (New et al., 2001).

	Lemme	Forme phonologique	Cumfreq [a]	Webcumfreq [b]
Adjectifs variables	*blanc*	*blɑ̃ / blɑ̃ʃ*	341,43	56,26
	brun	*brœ̃ / bryn*	52,74	67,95
	vert	*vɛʁ / vɛʁt*	127,19	30,64
Moyenne (ET)			173,79 (149,88)	31,23 (24,74)
Adjectifs invariables	*jaune*	*ʒon*	85,87	20,68
	rouge 'red'	*ruʒ*	219,03	31,38
	'black'	*nwaʁ*	364,7	41,20
Moyenne (ET)			223,29 (139,6)	31,09 (10,27)
Items de pratique	bleu	*blø*	170,49	25,48
	rose	*roz*	106,29	14,70

[a] Somme des fréquences orthographiques dans le corpus texte (New et al., 2001)
[b] Somme des fréquences orthographiques dans le web /1 000 (New et al., 2001)

ANNEXE B : Propriétés des adjectifs et noms utilisés pour les casse-têtes (adapté de Royle et Valois, 2010).

	Adjectifs			Noms		
	Variables	Invariables	Test-t	Féminins	Masculins	Test-t
Émergence	26,33 (1,53)	23,75 (2,50)	n.s.	22,83 (2,48)	20,50 (2,95)	n.s.
Phonèmes	3 (0)	3,25 (0,50)	n.s.	4,16 (1,33)	4,50 (0,84)	n.s.
Syllabes	1 (0)	1 (0)	_	1,50 (0,55)	1,83 (0,41)	n.s.
Fréquence lexicale	27,35 (25,37)	49,66 (31,19)	0,08	23,76 (24,50)	60,20 (93,04)	n.s.
Fréquence lemmatique	84,86 (52,25)	61,50 (53,35)	0,06	35,07 (42,80)	69.48 (96,02)	n.s.

Notes : Émergence = utilisé par 50% des enfants, en mois (Trudeau et al., 1999) ; Phonèmes = longueur en phonèmes ; Syllabes = longueur en syllabes (syllabation québécoise) ; Fréquence lexicale = fréquence orale (sous-titres des films, New et al., 2001) ; Fréquence lemmatique = fréquence orale (sous-titres des films, New et al., 2001)

ANNEXE C : Image du casse-tête pour l'induction d'adjectifs de couleur, version B



Évaluation des dysarthries et des apraxies de la parole : quelles approches ?

Autrice :
Michaëla Pernon[1,2]

Affiliations :
[1]Unité Parkinson-mouvements anormaux & CRMR Wilson, Hôpital Fondation Adolphe de Rothschild, Paris.

[2] Laboratoire de Phonétique et Phonologie, CNRS-Université Sorbonne Nouvelle, Paris.

Autrice de correspondance :
Michaëla Pernon
michaela.pernon@gmail.com

Dates :
Soumission : 19/11/2023
Acceptation : 03/11/2024
Publication : 28/12/2024

Comment citer cet article :
Pernon, M. (2024). Évaluation des dysarthries et des apraxies de la parole : quelles approches ? *Glossa*, 142, 29-48. https://doi.org/10.61989/6kkkds06

e-ISSN : 2117-7155

Licence :
© Copyright Michaëla Pernon, 2024
Ce travail est disponible sous licence Creative Commons Attribution 4.0 International.

Contexte. L'évaluation des troubles moteurs acquis de la parole (Motor Speech Disorders, MSDs), dysarthrie et/ou apraxie de la parole (Apraxia of Speech, AoS), s'effectue majoritairement en clinique dans le cadre d'une approche perceptive. Cette dernière demeure prédominante malgré les biais liés à la restauration phonétique/articulatoire ou à la mise en jeu d'une subjectivité relative aux représentations internes de l'auditeur. Par ailleurs, les principaux critères diagnostiques de l'AoS ne font pas, à ce jour, l'objet d'un consensus international. L'AoS partage, en outre, des signes avec les dysarthries qui souvent coexistent (Duffy, 2019). Dans ce contexte, les enjeux du diagnostic différentiel des MSDs restent majeurs.

Objectifs. Cet article présente une synthèse des études de notre thèse de doctorat. Celle-ci avait pour objectif d'examiner les apports des approches d'évaluation des dysarthries et des AoS, à savoir perceptive, acoustique et cognitive, pour le diagnostic et la caractérisation de ces MSDs.

Méthodes. Trois études ont été menées chez quatre groupes de locuteurs : neurotypiques, pathologiques présentant une AoS ou une dysarthrie hypokinétique dans la maladie de Parkinson ou une dysarthrie mixte flasque-spastique dans la sclérose latérale amyotrophique. Ces études, partiellement basées sur le corpus MonPaGe-2.0.s (Fougeron et al., 2018 ; Laganaro et al., 2021 ; Pernon et al., 2020) portaient respectivement sur chaque approche : (i) une tâche de classification auditivo-perceptive multijuges des MSDs, (ii) une analyse phonético-acoustique discrète par dimension de parole des locuteurs aux taux de précision les plus bas et les plus élevés en (i), (iii) un paradigme de double tâche évaluant, chez des locuteurs neurotypiques, les effets d'une demande attentionnelle et exécutive sur des tâches de parole et non verbales.

Résultats. Les approches d'évaluation perceptive et acoustique différaient peu pour le diagnostic des MSDs, ne permettant pas de diagnostiquer les MSDs très légers. L'approche acoustique a permis une caractérisation quantifiée des profils des MSDs en moyenne sur trois dimensions de parole évaluées par l'outil MonPaGe-2.0.s, comme dans l'étude de classification perceptive des MSDs. Les effets bidirectionnels de double tâche ont montré que le mode de présentation des stimuli, leur modalité, le type de tâche de parole, la demande attentionnelle des tâches non verbales, devaient être pris en compte dans l'élaboration et l'interprétation des performances. Ces trois études ont objectivé des différences selon les tâches de parole.

Conclusion. En complément de l'examen des dimensions de parole en jeu dans les tâches de parole, celui de leurs propriétés selon leur demande cognitivo-motrice (type de tâche de parole, complexité phonétique du corpus, condition de production), permet d'affiner le diagnostic différentiel des MSDs et d'améliorer leur caractérisation.

Mots-clés : dysarthrie ; apraxie de la parole ; diagnostic ; évaluation ; tâche de parole ; double tâche.

Assessment approaches for dysarthria and apraxia of speech.

Context. In clinical practice, acquired motor speech disorders (MSDs), dysarthria and/or apraxia of speech (AoS) are mainly assessed using a perceptual approach. This approach remains dominant despite the biases associated with phonemic restoration and the subjectivity of the listener's internal representations. Furthermore, there is a lack of international consensus on the main diagnostic criteria for AoS. Several of these features are shared with dysarthria, which also often coexists with AoS (Duffy, 2019). In this context, the assessment of MSDs for differential diagnosis is a major challenge.

Aims. This article provides an overview of the studies resulting from our PhD thesis. The aim of the thesis was to examine the contribution of the perceptual, acoustic and cognitive approaches to the assessment of dysarthria and AoS for the diagnosis and characterisation of these MSDs.

Methods. Three studies were conducted on 4 groups of speakers: neurotypical, pathological speakers with AoS, hypokinetic dysarthria in Parkinson's disease, and mixed flaccid-spastic dysarthria in amyotrophic lateral sclerosis. These studies, partly based on the MonPaGe-2.0.s corpus (Fougeron et al., 2018 ; Laganaro et al., 2021 ; Pernon et al., 2020), focused on each of the following assessment approaches: (i) a multi-judge auditory-perceptual classification task of MSDs, (ii) a discrete phonetic-acoustic analysis by speech dimension of the speakers with the lowest and highest accuracy rates in (i), (iii) a dual-task paradigm evaluating the effects of attentional and executive demands on speech and non-verbal tasks in neurotypical speakers.

Results. There was little difference between the perceptual and acoustic assessment approaches in diagnosing MSDs. The perceptual approach failed to diagnose very mild MSDs. The acoustic approach provided a quantified characterisation of MSDs profiles on average across three speech dimensions assessed by the MonPaGe-2.0.s tool, as in the study of perceptual classification of MSDs. The bidirectional dual-task effects showed that the mode of presentation of the stimuli and their modality, the type of speech task and the attentional demand of the non-verbal tasks had to be taken into account when developing and interpreting dual-task performance. These three studies found differences depending on the speech task.

Conclusion. In addition to considering the speech dimensions involved in speech tasks, a more precise examination of their cognitive-motor demands on speech (type of speech task, phonetic and linguistic complexity of the corpus, speech production condition) refines the differential diagnosis of MSDs and improves their characterisation during assessment.

Keywords: dysarthria; apraxia of speech; diagnosis; assessment; speech task; dual task.

INTRODUCTION

Cet article, rédigé dans le cadre du prix de thèse UNADREO 2023, présente une synthèse des études réalisées dans le cadre de notre thèse (Pernon, 2022b). Cette dernière a porté sur différentes approches d'évaluation des locuteurs présentant des troubles moteurs acquis de la parole (dysarthries et apraxies de la parole) : perceptive globale, phonético-acoustique discrète et cognitive, pour leur diagnostic et leur caractérisation.

Les troubles moteurs acquis de la parole

Les troubles moteurs de la parole, traduction de l'entité anglo-saxonne de « motor speech disorders » (MSDs) (Darley et al., 1975 ; Duffy, 2019), regroupent les dysarthries et les apraxies de la parole (AoS). Ces troubles sont secondaires à des lésions du système nerveux central et/ou périphérique. Pour l'AoS, elles concernent le plus souvent des aires et des réseaux cérébraux corticaux, à savoir : le gyrus frontal inférieur (pars opercularis), le cortex préfrontal dorso-latéral, le cortex prémoteur latéral supérieur, l'aire motrice supplémentaire et l'insula (Ballard et al., 2016 ; Duffy, 2019 ; Gorno-Tempini et al., 2011). Pour ce qui est des dysarthries, l'atteinte peut être aussi sous-corticale. De fait, toute atteinte d'un système neurophysiologique moteur est susceptible de les générer, à savoir les motoneurones central et périphérique, la voie cérébello-thalamo-corticale, le système extrapyramidal. Les MSDs peuvent par ailleurs être de différentes étiologies (vasculaire, neurodégénérative, traumatique, tumorale, infectieuse, métabolique, etc.). Ils représentent 56,7% des troubles de la communication orale d'étiologie neurologique, selon la cohorte de la Mayo Clinic située à Rochester (Minnesota, USA). Ils peuvent avoir des répercussions sur la communication quotidienne ainsi que sur la qualité de vie des locuteurs sur les versants émotionnels, sociaux et professionnels, en fonction de leur sévérité, du vécu du trouble et des besoins communicationnels des locuteurs (Duffy, 2019).

Les AoS ont été associées à une altération de la récupération et/ou de l'assemblage des plans moteurs de la parole (Code, 1998 ; Miller & Guenther, 2021 ; Van der Merwe, 1997, 2021 ; Ziegler, 2009), ou encore à une translation inefficace ou perturbée entre des représentations phonologiques intactes et des paramètres spatiotemporels appropriés nécessaires à la réalisation des mouvements articulatoires (McNeil et al., 2009). Les AoS ont également été attribuées à des troubles de la planification motrice dans le modèle de Van der Merwe (1997, 2021), niveau encore appelé programmation motrice dans d'autres modèles de contrôle moteur de la parole tels les modèles computationnels DIVA/GODIVA (Miller & Guenther, 2021 ; Tourville & Guenther, 2011).

Les dysarthries, quant à elles, sont définies comme un déficit du contrôle moteur et/ou de l'exécution des commandes neuromusculaires impliquées dans la production de la parole, secondaire à une atteinte du système nerveux central et/ou périphérique (Darley et al., 1975 ; Duffy, 2019 ; Van der Merwe, 1997, 2021). Six sous-types de dysarthries, déterminées selon le système neurophysiologique atteint sous-jacent, ont été initialement établis par l'équipe de Darley à la Mayo Clinic (Darley et al., 1975). Il s'agissait des dysarthries : flasque, spastique, ataxique, hypokinétique, hyperkinétique (choréique et dystonique) et mixte, auxquelles ont été ultérieurement ajoutées par Duffy (2005, dans Duffy, 2019) les dysarthries par atteinte unilatérale du motoneurone supérieur (Unilateral Upper Motor Neuron Dysarthria, UUMND) et les dysarthries d'étiologie indéterminée. Cette classification fait actuellement référence (Pernon, 2022a).

Approches d'évaluation des troubles moteurs acquis de la parole : enjeux pour leur diagnostic différentiel et leur caractérisation

La classification des dysarthries décrites ci-dessus et leur diagnostic, s'appuient sur des critères perceptifs après la prise en compte du/des système(s) neurophysiologique(s) atteint(s) des locuteurs (Darley et al., 1975 ; Kent, 1996). Pour ce qui est des critères diagnostiques proposés pour les AoS, ils s'avèrent également en grande partie perceptifs (Allison et al., 2020 ; Duffy, 2019 ; Jonkers et al., 2017 ; McNeil et al., 2009 ; Molloy & Jagoe, 2019 ; Strand et al., 2014). L'évaluation des locuteurs présentant ces MSDs se trouve alors principalement orientée en pratique clinique vers une approche perceptive immédiate et offrant une description rapide de la parole, actuellement considérée comme le « gold standard ». Par conséquent, cette dernière constitue la base pour le diagnostic, le jugement de sévérité, l'évaluation des changements dans la voix et la parole lors du suivi longitudinal et les décisions concernant la prise en soins des patients.

Cependant, à partir de cette approche perceptive globale, le diagnostic différentiel entre les MSDs

mais également avec la parole neurotypique, peut s'avérer parfois ardu. Une première source de biais est liée aux représentations internes des cliniciens/auditeurs, rendant la classification perceptive hautement subjective (Kim et al., 2011). Le phénomène de restauration phonémique peut également venir biaiser la perception, conduisant l'auditeur à remplacer les traits distinctifs d'un son déformé par un autre son et se confondre avec une paraphasie phonémique (Kent, 1996). À ces limites, viennent s'ajouter le partage de caractéristiques perceptives entre les dysarthries et les apraxies de la parole et leur co-existence fréquente. Selon la cohorte de la Mayo Clinic (Duffy, 2019), les dysarthries sont retrouvées dans 30% des cas lorsqu'une AoS (prédominante) est diagnostiquée. Enfin, l'absence - à ce jour - de consensus international sur les critères diagnostiques principaux des AoS, rend peu aisé le diagnostic différentiel entre ces deux MSDs.

Par ailleurs, cette évaluation perceptive prédominante, complétée plus récemment par des tests/batteries offrant une approche phonético-acoustique discrète, s'effectue classiquement sur la base d'un examen des différentes dimensions de parole. Elle ne propose généralement pas de déterminer objectivement les effets de la demande cognitivo-motrice de parole, ni de placer le locuteur avec un/des MSD(s) dans une/des situation(s) proche(s) de celles rencontrées en vie quotidienne, telle la condition de double tâche, consistant à parler et à effectuer autre chose simultanément (par exemple : écouter, marcher, conduire, etc.). En effet, les outils francophones d'évaluation de la parole disponibles, tels que la BECD (Auzou & Rolland-Monnoury, 2019), DIADOLAB (Sicard & Menin-Sicard, 2020a), VOCALAB (Menin-Sicard & Sicard, 2020 ; Sicard & Menin-Sicard, 2020b), ou encore MonPaGe-2.0.s (Fougeron et al., 2018 ; Laganaro et al., 2021 ; Pernon et al., 2020), n'examinent et/ou ne contrôlent pas ou peu cette demande cognitive exercée sur la parole en situation écologique, alors que cette dernière constitue davantage une règle qu'une exception au quotidien. Elle vient, de plus, modifier la demande attentionnelle et exécutive lors de la production motrice de la parole (Bailey & Dromey, 2015 ; Withfield et al., 2019).

Les enjeux de l'évaluation de la parole des locuteurs présentant un MSD restent ainsi majeurs à l'heure actuelle quant au choix des approches, des tâches de parole et des marqueurs/mesures retenus, permettant, entre autres, le diagnostic différentiel de ces troubles.

Au vu des biais liés à l'approche perceptive et des différents aspects rendant difficile la distinction des MSDs, nous pourrions supposer que le recours à des données phonétiques discrètes et acoustiques, en référence à des normes, pourrait faciliter le diagnostic différentiel des dysarthries et des apraxies de la parole ainsi que leur caractérisation. Par ailleurs, l'examen de la demande cognitivo-motrice de parole à travers la condition de double tâche, pourrait s'avérer complémentaire des approches précédentes, en prenant en compte la demande attentionnelle et exécutive en jeu dans la parole en vie quotidienne. Ici, sur la base de nos données issues d'études antérieures, il était question en premier lieu d'examiner, chez les locuteurs neurotypiques, les propriétés des tâches de parole et non verbales à prendre en compte pour l'élaboration ultérieure d'un outil destiné à l'évaluation des locuteurs présentant un trouble moteur acquis de la parole.

Ainsi, quelles approches d'évaluation de la parole seraient indiquées pour les troubles moteurs acquis de la parole à la fois pour leur diagnostic différentiel et leur caractérisation ?

Nous nous sommes plus précisément demandé (étude 1) d'une part si une approche auditivo-perceptive globale permettait de classer, de poser un diagnostic différentiel entre les dysarthries et les apraxies de la parole et plus spécifiquement pour quel degré de sévérité de MSD, quel degré d'expertise du professionnel et avec quelles tâches de parole.

D'autre part, ce diagnostic différentiel et cette caractérisation perceptifs seraient-ils améliorés par une approche phonético-acoustique discrète au moyen de marqueurs de parole, normés et validés, tels que proposés par l'outil MonPaGe-2.0.s (étude 2) ?

Enfin, quel pourrait être l'apport d'une approche cognitive de l'évaluation de la parole, ici explorée à travers l'effet de double tâche sur le débit de parole de locuteurs neurotypiques ? Quelles propriétés des tâches de parole faudrait-il alors retenir pour examiner cet effet (étude 3) ?

Les approches d'évaluation de la parole, perceptive globale (étude 1), phonético-acoustique discrète par dimension de parole (étude 2) chez des

locuteurs avec MSD, et cognitive portant sur la demande attentionnelle et exécutive de parole en double tâche (étude 3) chez des locuteurs neurotypiques, ont ainsi chacune fait l'objet d'une étude au sein de notre thèse (Pernon, 2022b).

Il est à noter que les deux premières reposaient principalement sur un corpus de parole issu de MonPaGe-2.0.s (Fougeron et al., 2018 ; Laganaro et al., 2021 ; Pernon et al., 2020), mentionné plus haut. Cet outil informatisé (Trouville et al., 2021) comporte 8 modules. La parole y est élicitée dans plusieurs types de tâches (lecture, répétition, série automatique, parole semi-spontanée, tâches de performances maximales : temps maximum de phonation (TMP) et diadococinésies verbales (DDK)) et sur du matériel linguistique de complexité variable. Ces modules couvrent, comme visé initialement, différentes dimensions de parole (voix, intelligibilité, précision articulatoire, prosodie, gestion temporelle) et plusieurs processus (planification-programmation-exécution). L'outil MonPaGe-2.0.s propose une approche perceptive discrète et acoustique de l'évaluation de la voix et de la parole, tout en prenant en compte les performances spécifiques de certaines tâches de parole. Suite à sa normalisation (*N*= 400 sujets francophones issus de Belgique, France, Québec et Suisse) et sa validation (*N*= 80 patients et 62 locuteurs contrôles neurotypiques différents de ceux de la base utilisée pour la constitution des normes), MonPaGe-2.0.s, destiné aux orthophonistes/logopèdes/logopédistes francophones, permet d'extraire une vingtaine de marqueurs perceptivo-acoustiques de voix et de parole de manière semi-automatique, à partir d'une cotation assistée sous *Praat*® (Boersma & Weenink, 2021). Il permet de diagnostiquer les MSDs légers à modérés-sévères, sans distinction à ce jour des AoS et des dysarthries, et d'en déterminer le degré de sévérité, avec une spécificité de 95% et une sensibilité de 83.8% (Laganaro et al., 2021). MonPaGe-2.0.s comporte deux applications « Passation » et « Cotation » (Trouville et al., 2021). sur la page dédiée à l'outil *MonPaGe* (2021) sur le site du Laboratoire de Phonétique et Phonologie, CNRS-Université Sorbonne Nouvelle, Paris.

Les études 1, 2 et 3 seront ici successivement résumées et leurs résultats discutés dans un dernier temps à travers leurs implications théoriques et cliniques. Il est à préciser que les sections suivantes constituent une vue d'ensemble et une synthèse des travaux réalisés dans le cadre de notre thèse de doctorat. Pour davantage de détails sur la méthode, les données traitées et les résultats, nous invitons le lecteur à se référer aux articles publiés dans le cadre de cette thèse, cités dans les sections suivantes, et/ou directement à notre thèse, accessible en ligne (Pernon, 2022b).

ETUDES

Étude 1. Classification auditivo-perceptive des troubles moteurs de la parole : rôle de la sévérité, de la tâche de parole et de l'expertise de l'auditeur.

Étant donné les limites du jugement perceptif global mentionnées dans l'introduction et le chevauchement des signes perceptifs de parole entre les sous-types de MSDs, quelques études ont été menées pour évaluer la fiabilité du système de classification des dysarthries de la Mayo Clinic (Bunton et al., 2007 ; Fonville et al., 2008 ; Lansford & Liss, 2014 ; Lansford et al., 2016 ; Van der Graaf et al., 2009 ; Zeplin & Kent, 1996 ; Zyski & Weisiger, 1987) ou encore pour valider des échelles d'évaluation perceptive des AoS (Diagnostic Instrument for Apraxia of Speech [DIAS] : Jonkers et al., 2017 ; Apraxia of Speech Rating Scale 1.0 [ASRS 1.0, dernière version : 3.5] : respectivement Strand et al., 2014 et Duffy et al., 2023). De rares travaux ont inclus à la fois des locuteurs avec AoS et des locuteurs dysarthriques (Hybbinette et al., 2021 ; Jonkers et al., 2017 ; Mumby et al., 2007). Les résultats de ces études ne semblent pas converger, cependant elles s'avèrent difficilement comparables en raison de plans d'expériences mettant en jeu des facteurs différents. Ces derniers affectent le processus décisionnel lors de la classification perceptive par les auditeurs, tels que les groupes de locuteurs (nombre de groupes, sous-types de MSD et pathologies sous-jacentes et inclusion d'un groupe de locuteurs neurotypiques), le degré de sévérité des MSDs, l'expertise/l'expérience clinique des auditeurs, les tâches de parole et le corpus (nombre et types de tâches de parole et durée des échantillons de parole), ou la méthode/la procédure expérimentale des expériences (tâches de classification à choix libre ou à choix forcé).

Compte tenu de ces éléments, nous avons élaboré une étude visant à examiner la précision du diagnostic différentiel des MSDs par des orthophonistes. Elle s'appuyait sur une tâche de classification auditivo-perceptive à choix

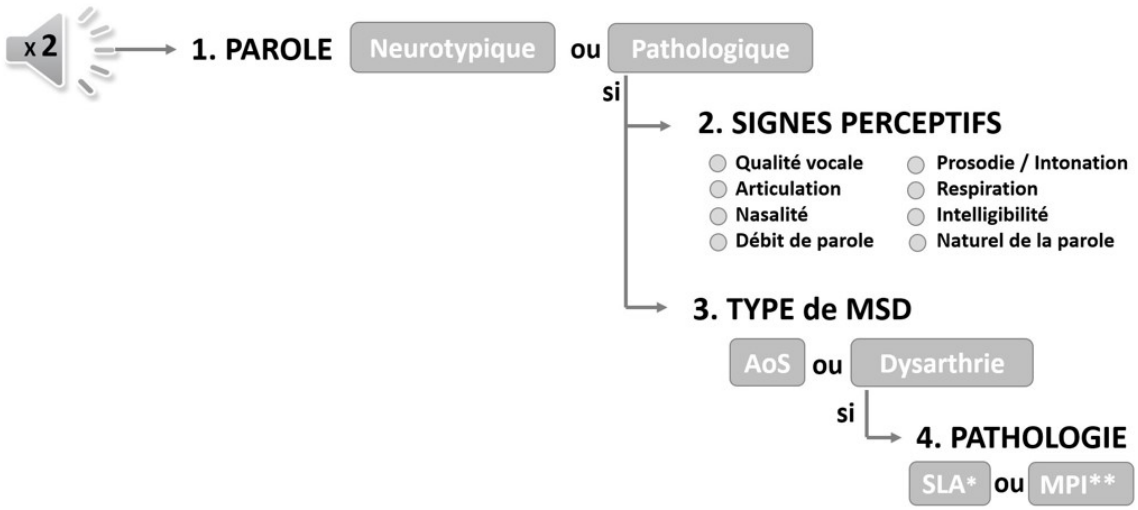


FIGURE 1 : Procédure expérimentale de l'expérience de classification auditivo-perceptive en ligne de l'étude 1.

forcé basée sur un raisonnement diagnostique conduisant, étape par étape, au diagnostic différentiel d'un sous-type de MSD à l'aide de 3 questions dichotomiques (Pernon, 2022b), procédure expérimentale susceptible de guider les cliniciens (adaptation de Duffy, 2019). Le lecteur pourra se référer au tableau 1 figurant plus bas pour une synthèse de cette étude.

Ce travail cherchait aussi à déterminer les variables ayant pu influencer la classification des locuteurs avec MSDs, à savoir la sévérité du MSD, l'expertise des auditeurs (étudiants en orthophonie et orthophonistes experts) et la tâche de parole (parole spontanée, texte lu à voix haute, DDK) issus du protocole MonPaGe-2.0.s (Fougeron et al., 2018 ; Laganaro et al., 2021 ; Pernon et al., 2020) suivant les consignes précisées dans le manuel (cf. page dédiée à l'outil MonPaGe (2021), site du Laboratoire de Phonétique et Phonologie, CNRS-Université Sorbonne Nouvelle, Paris).

À notre connaissance, ces variables n'ont jamais été prises en compte simultanément dans les études antérieures, ni testées dans une tâche de classification perceptive suivant un raisonnement diagnostique. L'identification de variables ayant un effet sur le diagnostic et le diagnostic différentiel des MSDs dans le cadre d'une approche auditivo-perceptive serait susceptible de nous permettre de mieux saisir les limites de cette approche et conduire à des recommandations pratiques pour l'améliorer. De plus, aucune tâche précédente de classification perceptive n'a impliqué un plan expérimental à choix forcé mélangeant des échantillons de parole de locuteurs présentant une AoS ou une dysarthrie avec ceux de locuteurs neurotypiques.

Enfin, nous avons également cherché à déterminer sur quelles caractéristiques perceptives les auditeurs s'appuyaient pour poser leur diagnostic.

L'article issu de cette étude (Pernon et al., 2022) est accessible en ligne.

Méthode (Étude 1)

Au sein de cette étude, ont été inclus 59 sujets, à savoir un groupe de locuteurs francophones neurotypiques (*N* = 29), des locuteurs francophones présentant un MSD léger à sévère (*N* = 30), dont un groupe de locuteurs avec AoS post-AVC prédominante (*N* =10) et deux groupes de locuteurs avec dysarthries : locuteurs avec dysarthrie hypokinétique (HD) dans le cadre de la maladie de Parkinson idiopathique (MPI) (*N* = 14) et locuteurs avec dysarthrie mixte flasque-spastique (MD-FlSp) (*N* = 6) dans la sclérose latérale amyotrophique (SLA) (*N* = 30). Les critères d'inclusion, les profils et les scores de sévérité du MSD de ces groupes locuteurs sont détaillés dans l'article en ligne en libre accès, cité précédemment.

Les extraits de parole étaient issus d'un plus large corpus de recherche (Fougeron et al., 2018 ; Laganaro et al., 2021 ; Pernon et al., 2020), se répartissaient en 3 tâches de parole issues de l'outil MonPaGe-2.0.s, incluant de la parole continue supposée refléter davantage le MSD et des tâches de performance, à savoir :

- des DDK dont les enregistrements ont été concaténés, comportant une série de DDK CV simples : « bababa » [bababa] suivie d'une série de DDK CCV complexes : « tratratra » [tRatRatRa], produites le plus rapidement et le plus précisément durant 4 s. ;

- deux paragraphes d'un texte lu à voix haute (concaténation des paragraphes « lundi » et « mercredi » du conte du protocole MonPaGe-2.0.s) ;

- un extrait de parole spontanée, réponse à la question : « Qu'avez-vous fait durant vos dernières vacances ou votre dernier weekend ? », de même durée que la lecture à voix haute.

Les extraits de parole ont été jugés par 40 auditeurs/juges orthophonistes, à savoir des orthophonistes experts (N = 20) et des étudiants en orthophonie en fin de Master 1 d'orthophonie (N = 20) qui avaient tous reçu un enseignement sur les MSDs et leur diagnostic différentiel.

Au sein de l'expérience de classification perceptive dont la procédure expérimentale est illustrée en figure 1, les échantillons de parole (N = 236) ont été divisés en 2 versions, de manière équilibrée, en fonction du sous-type de MSD, de sa sévérité, du genre et de l'âge des locuteurs.

L'expérience de classification auditivo-perceptive en ligne a été conçue sous la plateforme *Qualtrics* (2019), à partir de laquelle les juges l'ont également réalisée. Ils étaient informés au préalable des groupes de locuteurs et des sous-types de MSDs inclus.

Dans cette tâche auditivo-perceptive à choix forcé, les juges devaient répondre à des questions dichotomiques successives (de 1 à 4 question(s) par échantillon de parole) telles qu'elles sont illustrées figure 1, en suivant un raisonnement clinique et diagnostique les guidant pas à pas, tel qu'il est décrit par Duffy (2019) et adapté par Pernon (2022b).

Les analyses statistiques ont été effectuées sous R (*R Core Team*, 2019), à partir de modèles mixtes linéaires généralisés.

Résultats (Étude 1)

Le taux de précision globale de la classification effectuée par l'ensemble des juges a atteint une performance substantielle (72%), avec une bonne fiabilité inter-juges (coefficient de corrélation intraclasse : *ICC* = .78 [.44, .94]), mais elle variait en fonction de l'expertise des juges, de la tâche de parole et de la sévérité du MSD.

D'une part, les taux de précision se sont révélés supérieurs pour les échantillons de parole des locuteurs pathologiques (présentant un MSD) (86%) à ceux des locuteurs neurotypiques (57%) (question 1), pour les échantillons des locuteurs dysarthriques (83%) à ceux des locuteurs avec une AoS (59%) (question 2), et pour les échantillons des locuteurs présentant une HD (72%) à ceux avec une MD-FlSp (60%) (question 3). Dans l'ensemble, les locuteurs présentant un MSD plus sévère ont été les mieux classés.

D'autre part, les échantillons de parole continue (parole spontanée et texte lu à voix haute) ont obtenu des taux de classification supérieurs (moyennes de chaque bloc aux 3 questions dichotomiques : Parole spontanée : 74% ; Texte : 71% ; Texte + DDK : 69% ; DDK : 59%).

Un effet du degré d'expertise des juges a été également retrouvé avec un taux de classification plus élevé chez les orthophonistes experts (moyennes aux trois questions dichotomiques : 72%) que chez les étudiants en orthophonie (moyennes aux trois questions dichotomiques : 63%). La fiabilité inter-juges s'avérait également supérieure pour les orthophonistes experts (moyenne pour les 3 questions dichotomiques : orthophonistes experts : *ICC* = .81 [.56, .94] ; étudiants en orthophonie : *ICC* = .75 [.36, .95]).

Enfin, trois caractéristiques perceptives ont été utilisées en moyenne lors des classifications correctes. Leur type et leur combinaison différaient en fonction du MSD. Pour les locuteurs AoS, les taux les plus élevés concernaient les traits de débit de parole/fluence de la parole (25.02%), d'articulation (21.18%) et de naturel de la parole (15.88%). Pour les locuteurs avec une MD-FlSp dans le cadre d'une SLA, il s'agissait du débit de parole/fluence de la parole (16.90%), de la qualité vocale (16.21%) et de l'articulation (16.01%). Enfin, pour les locuteurs avec une HD dans le cadre d'une MPI, les taux les plus élevés correspondaient aux traits perceptifs de qualité vocale (23.47%), de débit de parole/fluence de la parole (15.84%) et enfin d'articulation (14.08%).

Les taux de précision obtenus dans cette étude (entre 58% et 86%), bien que majoritairement substantiels, se sont révélés supérieurs à la plupart de ceux des études de classification perceptive

précédentes (Bunton et al., 2007 ; Hybbinette et al., 2021 ; Jonkers et al., 2017 ; Mumby et al., 2007 ; Zyski & Weisiger, 1987). À partir des résultats de ces analyses perceptives effectuées sur des extraits de tâches de parole de MonPaGe.2.0.s, nous nous sommes interrogée sur l'apport d'une analyse phonético-acoustique discrète par dimension de parole pour la détection et la caractérisation de chacun des sous-types de MSDs évalués (AoS, HD, MD-FlSp). Nous avons ainsi examiné les profils MonPaGe-2.0.s des locuteurs correctement classés et incorrectement classés au sein de l'expérience de classification auditivo-perceptive ici présentée. Cela a fait l'objet de l'étude 2 présentée ci-après.

Étude 2. Approche phonético-acoustique discrète par dimension de parole dans l'évaluation des MSDs : analyses acoustiques et perceptives de la parole des locuteurs correctement et incorrectement classés dans l'expérience de classification perceptive, issues de MonPaGe-2.0.s.

L'étude 2 de notre thèse (Pernon, 2022b) a confronté l'approche perceptive globale du diagnostic différentiel des MSDs à une approche acoustique et perceptive, plus discrète (évaluation perceptive au niveau de la syllabe et/ou du phonème des modules intelligibilité et pseudomots) avec le logiciel MonPaGe-2.0.s. Elle a été effectuée sur les dimensions : intelligibilité de la parole, phonation/coordination pneumo-phonique, débit de parole/débit articulatoire, prosodie/intonation, précision articulatoire. Ces dimensions ont été testées à partir des 8 modules de MonPaGe.2.0.s afin d'examiner les apports de cet outil de screening dans le diagnostic et l'examen des MSDs.

Rappelons que MonPaGe-2.0.s (Fougeron et al., 2018 ; Laganaro et al., 2021 ; Pernon et al., 2020) a été développé afin de répondre au manque de marqueurs valides, sensibles et robustes en pratique clinique pour l'évaluation et le diagnostic des MSDs (Pommée et al., 2021), en offrant aux cliniciens un outil informatisé d'évaluation acoustique et perceptive, normé et validé, destiné aux locuteurs adultes francophones. Il ne permet pas, dans sa version actuelle, de distinguer les différents types et sous-types de MSDs selon les profils/marqueurs retrouvés, mais permet de déterminer si la parole/voix du locuteur évalué est pathologique.

Objectifs et méthode (Étude 2)

Cette étude visait tout d'abord à identifier quels étaient les locuteurs présentant un MSD qui avaient été incorrectement classés dans l'expérience de classification perceptive (groupe TPB) : locuteurs ayant reçu les Taux de Précision les plus Bas, inférieurs au taux de précision moyen obtenu par les locuteurs présentant le même MSD) et ceux correctement classés (groupe TPE) : locuteurs ayant reçu les Taux de Précision les plus Élevés, supérieurs au taux de précision moyen obtenu par les locuteurs présentant le même MSD) ayant fait l'objet de l'étude 1.

Les caractéristiques globales de ces deux groupes (démographiques, données cognitives, langagières, relatives à la motricité oro-faciale, à la voix/parole) ont été comparées.

Ont aussi été étudiés les marqueurs perceptifs et acoustiques/les scores de déviance de MonPaGe-2.0.s, qui caractérisaient les profils des locuteurs du groupe TPB par rapport au groupe TPE, par sous-type de MSD.

Il s'agissait enfin d'examiner les apports de l'outil de screening MonPaGe-2.0.s pour la pose du diagnostic de MSD chez les locuteurs du groupe TPB classés comme neurotypiques lors de la classification perceptive (étude 1).

Le protocole MonPaGe-2.0.s leur a été administré dans sa totalité. Les résultats perceptifs et acoustiques des locuteurs TPE et TPB, obtenus aux différents modules de l'outil MonPaGe-2.0.s, ont été comparés en fonction du type du MSD (groupes établis en fonction du MSD), ce, à partir des scores de déviance et des sous-scores de déviance fournis par le rapport MonPaGe-2.0.s.

Les résultats rapportés ici sont majoritairement descriptifs, qualitatifs, la taille des échantillons de locuteurs constitués en fonction du sous-type du MSD n'ayant pas permis la réalisation d'analyses statistiques comparatives.

Le chapitre 3 de notre thèse (Pernon, 2022b) pourra fournir au lecteur davantage de précisions sur la méthode de l'étude 2 et une présentation plus exhaustive de ces résultats.

Résultats (Étude 2)

Le groupe TPB comportait 9 locuteurs (2 avec AoS, 6 avec HD, 1 avec MD-FlSp), son taux de précision moyen en classification perceptive (étude 1) était

de 38% (*SD* = 14%). Le groupe TPE en comprenait 6 (2 avec AoS, 2 avec HD, 2 avec MD-FlSp) et son taux de précision moyen en classification perceptive (étude 1) s'élevait à 91% (*SD* = 3%).

Les données démographiques, cognitives, langagières et de motricité oro-faciale des groupes TPE et TPB ne différaient pas. Les degrés de sévérité des MSDs étaient significativement différents entre les deux groupes (définis à partir du score perceptif de la grille perceptive de la BECD (Auzou & Rolland-Monnoury, 2019) et du TotalDevScore de MonPaGe-2.0.s), le groupe TPE présentant un trouble plus sévère que le groupe TPB (TPE : score perceptif de la BECD : *M* = 9.5/20 ; *SD* = 3.8 ; TotalDevScore MonPaGe : *M* = 6.6/32 ; *SD* = 3.4) (TPB : score perceptif de la BECD : *M* = 4.6/20 ; *SD* = 4.3 ; TotalDevScore MonPaGe : *M* = 2.6/32 ; *SD* = 3.1).

Les locuteurs HD étaient surreprésentés parmi ceux incorrectement classés (groupe TPB), du fait sans doute du degré de sévérité très léger du MSD de la majorité d'entre eux (score perceptif de la BECD : *M* = 2.5/20 ; *SD* = 1.4 ; TotalDevScore MonPaGe : *M* = 1.6/32 ; *SD* = 2.5).

MonPaGe-2.0.s a permis de diagnostiquer le MSD de deux locuteurs avec HD considérés comme non pathologiques par les juges lors de la classification perceptive (étude 1). En revanche, des locuteurs avec une HD très légère (*N* = 4) n'ont pas non plus été diagnostiqués comme présentant un MSD au moyen de l'approche phonético-acoustique discrète proposée par MonPaGe-2.0.s. De plus, un locuteur du groupe TPB avec une MD-FlSp, considéré comme ayant une parole pathologique dans le cadre de la classification perceptive, n'a pas non plus été diagnostiqué comme porteur d'un MSD par MonPaGe.2.0.s. Ainsi, 5 locuteurs pathologiques n'ont pas été détectés par MonPaGe-2.0.s.

Enfin, l'outil MonPaGe-2-0.s a permis de diagnostiquer au total (diagnostic de parole pathologique > seuil pathologique) 1 locuteur supplémentaire par rapport à la classification perceptive.

À partir des données MonPaGe-2.0. des locuteurs TPB et TPE avec un MSD de sévérité légère-modérée, les profils suivants se sont dégagés :

- perturbations des locuteurs avec HD (*N* = 4) : intelligibilité de la parole > voix (f0[1] : jitter et écart-type f0) ;

- perturbations des locuteurs avec MD-FlSp (*N* = 3) : erreurs segmentales (données sur la précision articulatoire/module Pseudomots) = voix (f0 : jitter et écart-type f0) > DDK (débit AMR CV[2], AMR CCV[3], SMR CV[4]) ;

- perturbations des locuteurs avec AoS (*N* = 4) : erreurs segmentales > DDK (débit différence entre les SMR CV et AMR CV) > voix (CPPS[5]).

Les analyses descriptives et qualitatives des locuteurs TPB et TPE à partir des marqueurs phonético-acoustiques de MonPaGe-2.0.s retrouvent ainsi des déficits plus prototypiques dans le groupe TPE, aux MSDs plus sévères, et vont dans le sens des données de la littérature.

Le degré de sévérité du MSD, dont l'effet a été objectivé dans l'expérience perceptive, semble également au premier plan dans cette approche.

Les marqueurs pathologiques dégagés via les scores de déviance de MonPaGe.2.0.s pour chaque type et sous-type de MSD concordaient en grande partie avec les traits perceptifs utilisés par les juges lors de la classification perceptive. Toutefois les tâches de parole, comme les DDK et les marqueurs vocaux, qui étaient peu contributives en classification perceptive, se sont révélées fournir ici des indices acoustiques pertinents, caractéristiques des patterns spécifiques des types de MSD. De fait, il a été montré chez les locuteurs avec AoS que le débit et la précision articulatoires en DDK séquentielles étaient significativement inférieurs à ceux des DDK répétitives, dissociation non retrouvée chez les locuteurs dysarthriques (Allisson et al., 2020 ; Duffy, 2019 ; Lancheros et al., 2022). Dans la classification perceptive, seules des DDK CV et CCV répétitives avaient été incluses, ce qui ne permettait pas aux juges de s'appuyer sur cet effet de tâche de parole. Les sous-scores de déviance du module DDK de MonPaGe-2-0.s intègrent ces différents types de DDK ainsi

1 f0 : fréquence fondamentale.

2 AMR CV : diadococinésies verbales répétitives Consonne – Voyelle, telle [ba ba ba].

3 AMR CCV : diadococinésies verbales répétitives Consonne – Consonne- Voyelle, telle [tRa tRa tRa].

4 SMR CV : diadococinésies verbales séquentielles Consonne-Voyelle, telle [badego].

5 CPPS : Smoothed cepstral peak prominence.

qu'un marqueur reflétant cette dissociation (DDK SMR – DDK AMR), permettant ainsi de mettre en évidence cet effet, notamment dans le groupe TPE ; ils pourraient être utilisés dans une étude subséquente.

Les approches perceptive globale et phonético-acoustique discrète par dimension de parole s'avèrent ainsi complémentaires pour la pose du diagnostic de MSD et sa caractérisation. L'approche phonético-acoustique discrète par dimension de parole a par ailleurs permis d'examiner plus objectivement les dimensions de voix et de gestion temporelle de la parole. Bien que la généralisation des résultats de cette étude demeure limitée compte tenu de la faible taille des sous-groupes de locuteurs, cette étude fournit des premières pistes sur les profils de MSDs par dimension de parole et sur le diagnostic différentiel entre MSDs via MonPaGe-2.0.s.

Étude 3. Approche cognitive de l'évaluation de la parole : étude exploratoire de l'effet bidirectionnel de double tâche chez des locuteurs neurotypiques en fonction de la nature et des propriétés des tâches de parole et non verbales.

Cette étude (Pernon et al., 2019) aborde l'approche cognitive de l'évaluation de la parole à partir de l'étude de la demande attentionnelle et exécutive de la parole en double tâche chez des locuteurs neurotypiques. Cette dernière est généralement étudiée dans le cadre d'un paradigme de double tâche qui consiste à comparer les performances à une tâche réalisée en condition isolée à celles obtenues à cette même tâche effectuée simultanément à une autre GHDtâche (condition de double tâche). Cette approche, peu explorée en clinique, est plus expérimentale que celles des deux études précédentes.

L'implication de la demande attentionnelle et exécutive sur les niveaux et processus de production motrice de la parole, longtemps débattue, est de fait aujourd'hui plus largement admise. Elle est mentionnée par Van den Merwe dans ses deux modèles (1997, 2021) et par Hickok (2014) dans son modèle HSFC (Hierarchical State Feedback Control Model). Les données de neuroimagerie ont montré, quant à elles, des interactions entre la production motrice de parole et l'attention et les fonctions exécutives/contrôle inhibiteur, partagées par les mêmes réseaux cérébraux (Dick et al., 2019). À la lumière de ces données, il semblerait exister une/des forme(s) d'attention intrinsèque à la production motrice de la parole, « encapsulée » plus ou moins disponible selon différents facteurs (Van der Merwe, 1997) et une/des autre(s) forme(s) d'attention externe, qui agirai(en)t, tel un monitoring, sur certains de ces niveaux et/ou processus (Hickok, 2014 ; Van der Merwe, 1997).

L'étude présentée dans ce chapitre fait suite à de premiers travaux exploratoires qui avaient porté sur l'impact de la double tâche sur la gestion temporelle de la parole de locuteurs dysarthriques atteints de maladie de Wilson et de locuteurs neurotypiques appariés.

L'étude préliminaire 1 (Pernon et al., 2013) avait examiné l'effet de double tâche sur le débit de parole en répétition d'un énoncé lors de la réalisation d'une tâche visuo-motrice au mode continu de présentation des stimuli (tâches de barrage papier crayon). Une accélération du débit de parole avait été observée chez les locuteurs contrôles neurotypiques et chez une partie des locuteurs dysarthriques wilsoniens.

L'étude préliminaire 2 (Pernon et al., 2017 ; Pernon, 2018) avait montré, quant à elle, un ralentissement du débit articulatoire en double tâche lors de la production de la série automatique des jours de la semaine réalisée simultanément à une tâche visuo-spatiale informatisée de mode discret de présentation des stimuli (appui sur une touche du clavier lors de l'apparition de stimuli à l'écran).

Dans ces deux études, de nombreux facteurs entraient en jeu dans l'interférence observée pour des tâches réalisées simultanément. De fait, il a été montré que la modalité (Wickens, 2008) et le mode de présentation des stimuli à traiter (Pashler et al., 2001) pouvaient affecter le chevauchement dans le temps des deux tâches (Eichhorn et al., 2018). De plus, la priorité donnée à une tâche par rapport à l'autre et la vitesse allouée à leur réalisation viennent aussi moduler les effets de double tâche (Fournet et al., 2021 ; Pashler et al., 2001). Enfin, la prise en compte du degré d'automaticité des tâches, de leur coût cognitif et de leur complexité s'avère cruciale pour l'étude de l'effet de double tâche (Pashler et al., 2001).

L'objectif de l'étude présentée dans ce chapitre était de chercher à mieux saisir, en se centrant sur les locuteurs neurotypiques, les résultats divergents des deux études préliminaires, dont la nature ainsi que les propriétés des tâches de parole et non verbales différaient.

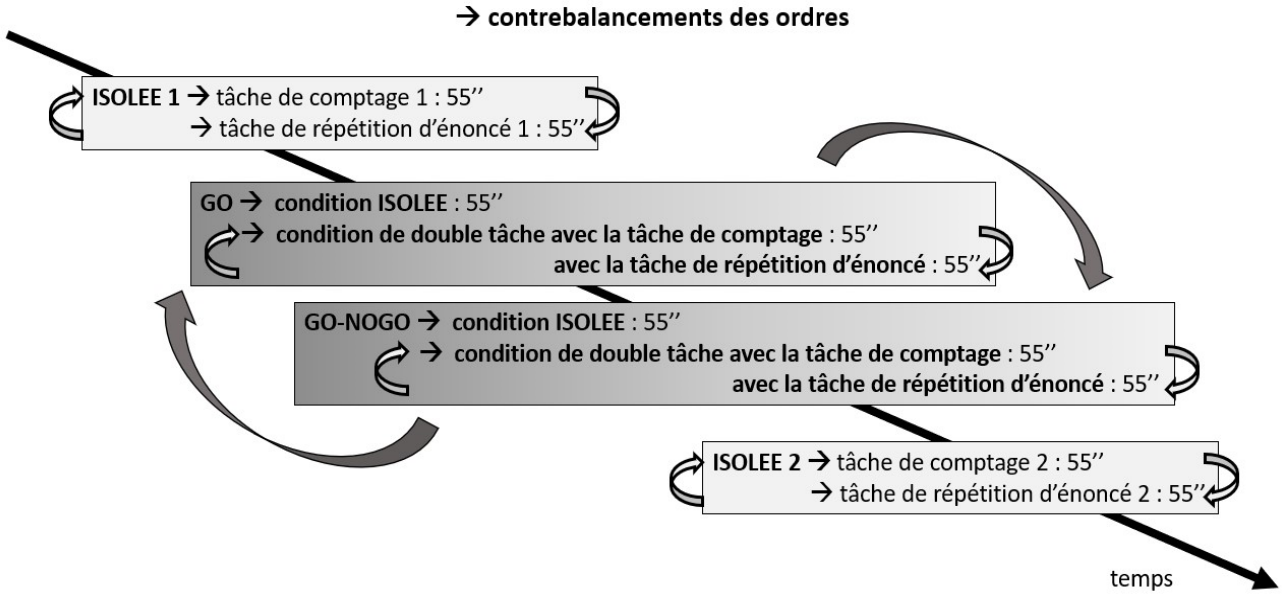


FIGURE 2 : Procédure expérimentale des expériences 1 et 2 de l'étude 3.

Nous avons ainsi souhaité déterminer s'il existait un effet de double tâche sur la parole et sur les performances en tâches non verbales dans le cadre d'une analyse bidirectionnelle chez des locuteurs neurotypiques. Ont ainsi été pris en compte les résultats obtenus à la fois aux tâches de parole (débit de parole) et aux tâches non verbales (temps de réaction, taux de réponses correctes, nombre de stimuli correctement barrés).

Ce travail visait également à examiner l'influence des propriétés des tâches (degré de demande attentionnelle des tâches non verbales, type de tâche de parole, mode de présentation des stimuli) sur l'effet de double tâche.

Enfin, à partir des résultats des analyses bidirectionnelles, il s'agissait d'étudier les stratégies attentionnelles utilisées par les locuteurs neurotypiques pour la réalisation de la double tâche.

Méthode (Étude 3)

Pour ce faire, nous avons inclus dans cette étude 27 locuteurs neurotypiques et combiné les facteurs précédemment mentionnés au sein de la procédure expérimentale de deux nouvelles expériences réalisées par les mêmes locuteurs neurotypiques dans le cadre de paradigmes de double tâche.

Le mode de présentation des stimuli des deux tâches non verbales de l'une était discret (expérience 1) et celui de l'autre continu (expérience 2). Les tâches non verbales des études précédentes (1 tâche de réaction simple : go et 1 tâche de go-nogo) ont été reprises et homogénéisées pour les besoins de cette nouvelle étude. Les deux types de tâches de parole (série automatique - ici comptage de 1 à 20 - et répétition d'un énoncé) ont été intégrés et les analyses de parole se sont focalisées uniquement sur le paramètre temporel du débit de parole. Chacune des tâches de parole enregistrées sous *Audacity*® (version 2.1.0, 2015), et chacune des tâches non verbales, d'une durée de 55", ont été administrées aux sujets en condition isolée, puis dans les deux conditions de double tâche.

Enfin, nous avons veillé à corriger les biais relevés dans les deux précédentes études, à savoir contrebalancer les tâches, modifier les consignes relatives à la vitesse de traitement (« le plus vite possible »), supprimées ici, et prendre en compte les éventuels effets d'apprentissage ou de fatigabilité entre les conditions pour les tâches de parole en ajoutant une condition isolée finale, dont les résultats ont été moyennés avec la tâche de parole isolée initiale. La procédure expérimentale des expériences 1 et 2 est illustré en figure 2.

Résultats (Étude 3)

La modification du débit de parole en double tâche a été interprétée comme un indice de la demande attentionnelle requise par la production de parole. L'analyse bidirectionnelle des effets d'une tâche

sur l'autre constituait un moyen d'estimer les mécanismes sous-jacents et les stratégies globales adoptées par les locuteurs lors d'une double tâche.

Nos résultats ont tout d'abord montré que les effets de double tâche n'étaient pas similaires d'une tâche de parole à l'autre. Un effet de double tâche sur le débit de parole, ralenti, a été trouvé pour la tâche de comptage dans les deux expériences. En revanche, nous n'avons pas mis en évidence d'effet de double tâche sur le débit de parole en répétition d'énoncé.

Le traitement des tâches non verbales était également modifié lors de la parole, mais la direction de l'effet dépendait du type de tâche de parole et non verbale. Dans l'expérience 1 (mode discret des tâches non verbales) était relevé un effet de double tâche négatif marqué par un allongement des temps de réaction des réponses correctes en tâches de go et de go-nogo et un ralentissement du débit de parole en tâche de comptage lors du go-nogo. Dans l'expérience 2 (mode continu des tâches non verbales), un effet de double tâche positif était mis en évidence avec une augmentation du nombre de stimuli correctement traités lors de la répétition d'énoncé où le débit de parole tendait à s'accélérer (et non lors du comptage).

Ainsi, ces résultats sont venus confirmer le recrutement de ressources attentionnelles dans la production motrice de la parole, ce, en fonction des propriétés des tâches. Les mécanismes ou stratégies sous-jacents utilisés par les sujets sains pour réaliser la double tâche dépendent du type de tâche de parole, de la demande attentionnelle des tâches non verbales, du mode de présentation des stimuli, mais également de leur modalité.

La théorie des capacités partagées (« capacity sharing theory ») (Kahneman, 1973) permet d'expliquer la stratégie adoptée dans l'expérience 1 (diminution des performances dans les deux tâches). Dans l'expérience 2, la tendance à l'accélération du débit de parole en répétition d'énoncé et l'augmentation du nombre de stimuli correctement traités pourraient refléter un "effet d'aimant" (« magnet effect »), d'attracteur, relatif à la tendance des oscillateurs biologiques à s'attirer mutuellement (Ebersbach et al., 1995), favorisée par la modalité motrice commune de traitement des deux tâches à l'origine d'un « effet énergisant » (« energizing effect ») (Adams et al., 2010).

DISCUSSION GÉNÉRALE ET CONCLUSIONS

Notre thèse avait pour objectif d'examiner les apports respectifs des approches perceptive globale, phonético-acoustique discrète par dimension de parole et cognitive à travers un paradigme de double tâche dans l'évaluation de la parole de locuteurs présentant un/des trouble(s) moteur(s) acquis de la parole, dysarthrie et/ou apraxie de la parole.

Nous discuterons à présent des résultats des études 1, 2 et 3 présentés plus haut, à travers leurs implications théoriques et cliniques (Pernon, 2022b). Une synthèse de leurs apports est proposée dans le tableau 1.

Implications théoriques

Les approches perceptive globale et phonético-acoustique discrète se sont révélées complémentaires dans l'évaluation de la parole des locuteurs présentant un MSD, fonction du type de MSD et de sa sévérité, des objectifs fixés pour l'examen de la parole.

Elles étaient quasi-équivalentes dans une visée diagnostique de l'évaluation des MSDs de formes légère à modérée. Ce diagnostic était établi en moyenne sur 3 signes perceptifs dans l'approche perceptive globale et sur 3 scores de déviance, tels qu'ils sont détaillés dans la section Résultats des études 1 et 2, au moyen de l'outil de screening MonPaGe-2.0.s.

Pour la caractérisation des profils dysarthriques, l'approche phonético-acoustique discrète par dimension de parole vient compléter l'approche perceptive globale par une description plus détaillée, quantifiée et objective de la parole des locuteurs présentant un MSD. Les marqueurs acoustiques vocaux et temporels, perceptifs discrets relatifs à l'articulation et à l'intelligibilité, ainsi que ceux en lien avec les tâches de parole DDK se sont révélés les plus sensibles.

Le paramètre de sévérité vient jouer un rôle sur l'expression des signes respectifs des MSDs qui peuvent s'en trouver fortement modifiés et fausser leur diagnostic. Pour les locuteurs avec HD très légères, le caractère pathologique de la parole est interrogé car leur dysarthrie n'a été détectée ni par l'approche perceptive globale, ni par l'approche phonético-acoustique discrète via MonPaGe-2-0.s. Des analyses sur de plus larges corpus de parole ou la contribution d'autres marqueurs de parole

TABLEAU 1 : Résumé des principaux résultats et apports des trois études sur les approches d'évaluation perceptive globale, phonético-acoustique discrète par dimension de parole et cognitive exploratoire (demande attentionnelle et exécutive en double tâche).

APPROCHES	PARTICIPANTS (N)	MESURES	RÉSULTATS	APPORTS (diagnostic, caractérisation, stratégies)
PERCEPTIVE **GLOBALE** (classification perceptive)	**Locuteurs : N = 59** **Neurotypiques** : **N = 29** **Pathologiques : N = 30** (AoS : N = 10 HD : N = 14 MD-FlSp : N = 6) **Auditeurs : N = 29** Orthophonistes : N = 20 Étudiants : N = 20	- taux de précision - analyse des erreurs - fiabilité inter-juges (ICC) - nb et type de traits perceptifs	- **Effet du type / sous-type de MSD** : - neurotypique > pathologique - dysarthrie > AoS - HD > MD-FlSp **- Effet du degré de sévérité du MSD** : - modéré-léger > léger-très léger **- Effet de l'expertise des auditeurs** : - orthophonistes > étudiants **- Effet de la tâche de parole** : - Spontané, Texte > Texte+DDK, DDK - Fiabilité inter-juges : bonne - Nb moyen de traits perceptifs : 3 - Analyses des erreurs de classification : - AoS → MD-FlSp - HD → Neurotypique - MD-FlSp → HD	- Taux de précision : substantiels > études antérieures - Efficacité du design reposant sur un raisonnement clinique pas à pas pour le diagnostic différentiel des MSDs → arbre décisionnel. - Diagnostic différentiel plus aisé pour : - les dysarthries légères-modérées - les tâches de parole continue - les orthophonistes experts
PHONÉTICO-ACOUSTIQUE **DISCRÈTE** **PAR DIMENSION DE PAROLE** (analyses et profils MonPaGe-2.0.s) - sous-scores de déviance	**Locuteurs du groupe TPB : N = 9** (6 HD, 2 AoS, 1 MD-FlSp) **Locuteurs du groupe TPE : N = 6** (2 HD, 2 AoS, 2 MD-FlSp)	- score total e-GeBAS - score total ASRS 1.0 - score perceptif BECD - score capacités motrice oro-faciale BECD - score total MMSE **MonPaGe-2-0.s :** - TotalDevScore - scores de déviance	**- Effet de sévérité** : TPB < TPE - Nb de locuteurs HD : TPB > TPE - 4 locuteurs HD très légers non diagnostiqués - 1 locuteur MD-FlSp très léger non diagnostiqué - Différence diagnostique entre perceptif global et phonético-acoustique : + 1 locuteur - Nb moyen de scores de déviance : 3 **Profils MonPaGe.2.0.s par type de MSD :** - **HD** : intelligibilité de la parole > voix (f0 : jitter et ET f0) **- MD-FlSp** : erreurs segmentales = voix (f0 : jitter et ET f0) > DDK (débit AMR CV, AMR CCV, SMR CV) **- AoS** : erreurs segmentales > DDK (débit différence SMR CV - AMR CV) > voix (CPPS).	- Pour ces locuteurs : diagnostic de MSDs : perceptif global ≈ phonético-acoustique (+1) - Caractérisation possible des signes des MSDs au moyen de l'outil MonPaGe.2.0.s selon : - dimensions de parole / intelligibilité - tâches de parole : DDK –> contributives - paramètres acoustiques vocaux - Pose du diagnostic de MSDs très légers → sensibilité insuffisante de MonPaGe.2.0.s
COGNTIVE EXPLORATOIRE **CENTRÉE SUR** **LA DEMANDE ATTENTIONNELLE** **ET EXÉCUTIVE DE PAROLE** (conditions de double tâche)	**Locuteurs neurotypiques : N = 27**	**Parole :** - nb de syll./sec. **Tâches non verbales :** **Expérience 1 (mode discret)** - temps de réaction - taux de réponses correctes **Expérience 2 (mode continu)** - nb de stimuli correctement traités	**- Effets de double tâche fonction du mode de présentation des stimuli :** - tâches de parole : ralentissement en comptage - tâches non verbales : ralentissement des tâches discrètes et accélération des tâches continues **- Effets du degré de demande attentionnelle** - tâches GO-NOGO (2 modes), comptage : inhibition ++ **- Effets bidirectionnels** - comptage et tâches non verbales	- Mise en évidence d'une demande attentionnelle et exécutive de la production motrice de la parole chez des locuteurs neurotypiques → débit de parole : programmation / exécution motrice ? - Analyse bidirectionnelle des effets de double tâche → stratégies adoptées : - capacity sharing : comptage/tâches non verbales - magnet /energizing effects : Répétition d'un énoncé /GO-NOGO continu - DBT:prise en compte des propriétés des tâches

Notes. AoS : apraxie de la parole ; HD : dysarthrie hypokinétique ; MD-FlSp : dysarthrie mixte flasque-spastique ; MSD : trouble moteur acquis de la parole ; f0 : fréquence fondamentale ; DDK AMR CV : diadococinésies verbales répétitives Consonne-Voyelle, telle [ba ba ba] ; DDK AMR CCV : diadococinésies verbales répétitives Consonne-Consonne- Voyelle, telle [tRa tRa tRa] ; SMR CV : diadococinésies verbales séquentielles Consonne-Voyelle, telle [badego] ; MMSE : Mini-Mental State Examination, version GRECO (Kalafat et al., 2003) ; e-GeBAS : Electronic Geneva Bedside Aphasia Scale (Chicherio et al., 2019) ; CCPS : smoothed cepstral peak prominence ; ASRS 1.0 : Apraxia of Speech Rating Scale 1.0 (Strand et al., 2014) ; BECD : Batterie d'Évaluation Clinique des Dysarthries (Auzou & Rolland-Monnoury, 2019) ; TPB : groupe de locuteurs pathologiques ayant reçu les taux les plus bas à l'expérience de classification perceptive ; TPE : groupe de locuteurs pathologiques ayant reçu les taux les plus hauts à l'expérience de classification perceptive.

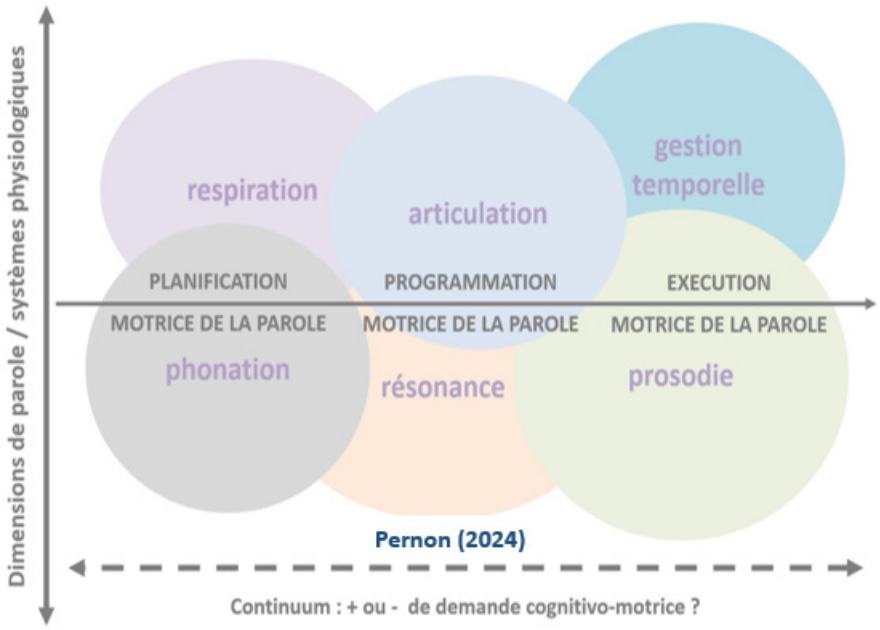


FIGURE 3 : Propriétés des tâches de parole : dimensions de parole en fonction de la demande cognitivo-motrice de parole.

auraient peut-être permis de détecter leur MSD, comme ceux relatifs à l'intonation par exemple, encore peu développés dans la version actuelle de MonPaGe-2.0.s. Il pourrait être ainsi envisagé de spécifier pour un même MSD des signes/critères distinctifs, des profils selon sa sévérité (très léger, léger, modéré, sévère), voire de les intégrer à leur classification. La validité d'une telle proposition reste cependant discutable compte tenu de l'importante variabilité interindividuelle (et intra-individuelle) des présentations des locuteurs d'un même MSD et pour un même degré de sévérité.

Par ailleurs, pour les locuteurs présentant une AoS, dont les taux de précision en classification perceptive globale étaient insuffisants, les données issues de l'approche acoustico-phonétique pourraient venir faciliter la pose du diagnostic par la mise en évidence de certains signes pathognomoniques (Melle & Gallego, 2012 ; Ziegler, 2002). Parmi ceux-ci figurent la dissociation retrouvée pour le débit articulatoire chez les locuteurs AoS du groupe TPE : celui obtenu lors de la tâche de DDK SMR CV étant inférieur à ceux obtenus lors des tâches de DDK AMR CV (Lancheros et al., 2022).

L'examen de la demande attentionnelle et exécutive lors de la production motrice de parole en double tâche chez les neurotypiques au sein de l'étude 3 a montré la nécessité de prendre en compte les propriétés des tâches de parole et non verbales, telles que le mode de présentation de leurs stimuli, afin de pouvoir interpréter précisément les effets de double tâche. Le paradigme de double tâche pourrait avoir sa place dans le cadre d'une approche cognitive de l'évaluation de la parole, offrant des conditions de parole plus proches de celles rencontrées en vie quotidienne. Chez les locuteurs pathologiques, les effets de double tâche sur la parole ont été majoritairement étudiés sur des paramètres de gestion temporelle de la parole (Fournet et al., 2021 ; Ho et al., 2002 ; Kemper et al., 2009, 2011 ; Whitfield & Goberman, 2017), sur lesquels était également centrée notre étude. Ils ont été retrouvés chez des locuteurs présentant un MSD, notamment chez ceux avec une HD liée à une maladie de Parkinson (Dromey & Bates, 2005 ; Ho et al., 2002 ; Whitfield et al., 2019). Il serait pertinent d'analyser d'autres paramètres de parole, potentiellement plus sensibles à l'effet de la double tâche. En ce qui concerne les locuteurs AoS, des travaux supplémentaires seraient nécessaires. En effet, les rares études dont nous disposons montrent peu d'effets et de différences significatives avec les locuteurs neurotypiques sur les paramètres étudiés (Harmon et al., 2019 ; Southwood & Dagenais, 2001). D'autres paramètres segmentaux, suprasegmentaux, vocaux, plus spécifiques de la production motrice/du contrôle moteur de la parole que les paramètres temporels de parole, ont fait l'objet d'études (chez des locuteurs parkinsoniens dysarthriques : Withfield et al., 2019 : effets de la variation de f0 ; Ho et al., 2002 :

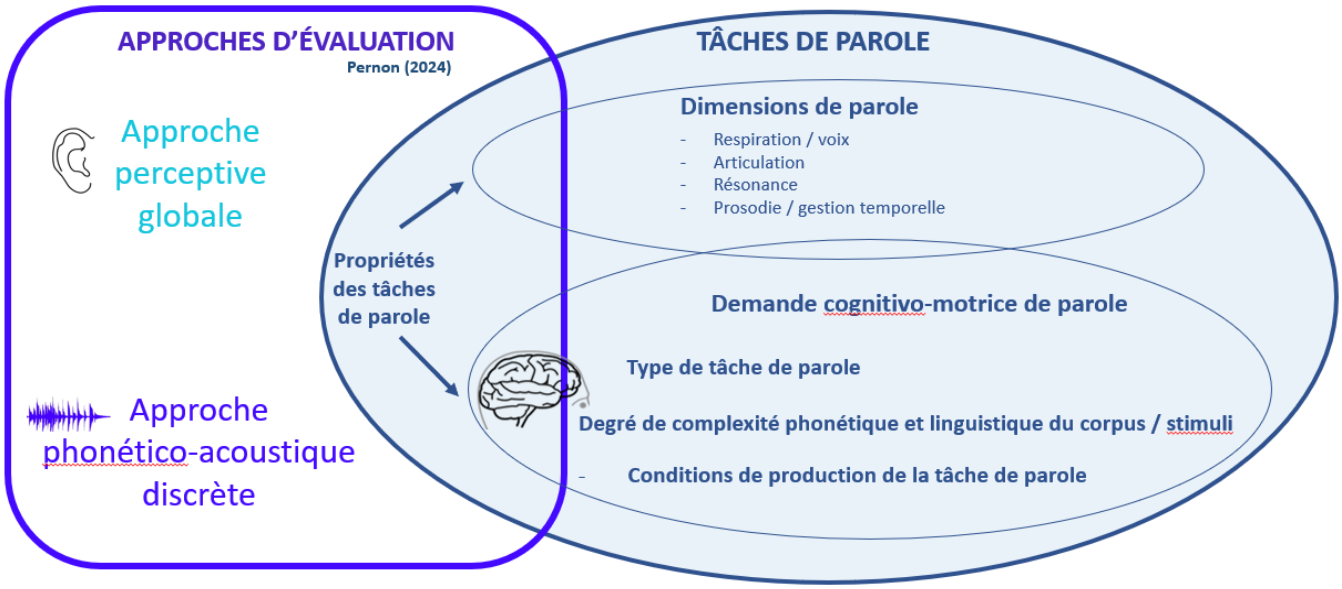


FIGURE 4 : Approche cognitive transversale de l'évaluation de la parole à travers les propriétés des tâches de parole.

réduction progressive de l'intensité vocale ; chez des locuteurs avec AoS : Southwood & Dagenais, 2001 : augmentation des allongements de sons et des distorsions). Ces paramètres pourraient constituer de potentiels marqueurs pour mesurer les effets de la double tâche sur la production motrice de la parole chez les locuteurs avec MSD.

Ces études portant sur des locuteurs pathologiques n'intègrent cependant qu'un seul type de MSD. Des études comparatives des effets de double tâche sur la parole de locuteurs présentant différents types de MSDs seraient ainsi nécessaires afin d'en examiner plus avant l'intérêt diagnostique. Ce dernier pourrait être établi à travers la recherche d'une modulation de la demande attentionnelle et exécutive aux différents niveaux de production motrice de la parole, à savoir la planification, la programmation et l'exécution motrices selon le FL model de Van der Merwe (1997, 2021). Ce(s) niveau(x) et/ou processus de production motrice de parole serai(en)t alors déterminé(s) par les paramètres de parole affectés par l'interférence attentionnelle en double tâche. Le traitement/contrôle de ces paramètres a de fait été rattaché à certains niveaux/processus de production motrice de la parole (Van der Merwe, 2021).

Les propriétés des tâches de parole pourraient être davantage spécifiées, à savoir les dimensions de parole mises en jeu et leurs interactions avec la demande cognitivo-motrice de parole, telles qu'illustrées au sein de la figure 3.

Les effets de certains paramètres de la demande cognitivo-motrice de parole, appelés « facteurs contextuels » par Van der Merwe (1997), ont, de fait, été retrouvés. La caractérisation plus systématique et plus précise des propriétés des tâches de parole pourrait s'effectuer en indiquant, par exemple, + ou - de degré d'automaticité, de degré de familiarité, de degré de complexité phonétique (structure syllabique et nombre de syllabes produites, etc.). L'ensemble de ces données (approches et tâches de parole) viendront en retour nourrir les connaissances et préciser les caractéristiques, les signes spécifiques de chaque MSD et leur diagnostic différentiel.

Implications cliniques

Une structuration de l'évaluation de la parole des locuteurs présentant un MSD pourrait être proposée, ce à partir des composantes : approches, outils normés et marqueurs, prenant en compte les propriétés des tâches de parole (dimension de parole en jeu, demande cognitivo-motrice). Elles concernent la nature de la tâche de parole (degré de familiarité, degré d'automaticité, etc.), la complexité articulatoire du corpus mis en jeu par des variables psycholinguistiques/phonétiques et les conditions de production de la tâche de parole.

L'examen de cette demande cognitivo-motrice des tâches de parole permettrait d'affiner le diagnostic différentiel des MSDs et leur caractérisation et constituerait l'approche cognitive de l'évaluation de la parole à proprement parler. La figure 4 vient schématiser ce propos. Il est à noter que nous n'abordons ici ni l'approche motrice-articulatoire, ni l'approche émotionnelle. L'apport de l'autoévaluation de sa parole et de sa qualité

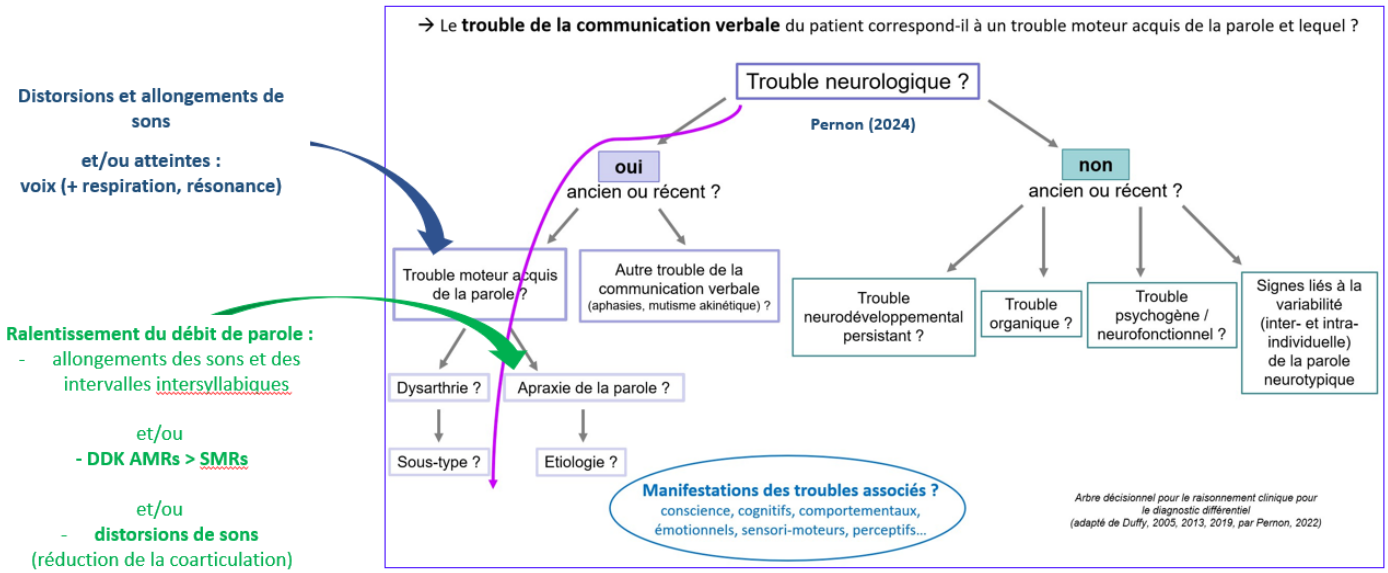


FIGURE 5 : Arbre décisionnel pour le diagnostic différentiel des MSDs.

de vie par le locuteur et celui, complémentaire, des données biologiques, génétiques, neurophysiologiques ou encore de neuroimagerie, ne sont pas non plus représentés, nos études n'ayant pas porté sur ces aspects.

Le recours à un raisonnement guidé pour le diagnostic différentiel de MSDs pourrait être par ailleurs encouragé.

Les résultats de l'expérience de classification perceptive des MSDs (étude 1) (Pernon et al., 2022) qui s'appuyait sur des questions dichotomiques successives, ont montré que le recours à un raisonnement diagnostique clinique pas à pas s'inspirant de celui développé par Duffy (2019) était facilitateur. Ce type de procédure expérimentale a probablement influencé la précision des diagnostics, dont les taux étaient plus élevés que ceux des travaux antérieurs de classification qui ne reposaient pas sur les mêmes types (classification libre : Zyski & Weisiger, 1987 ; liste fermée comportant plusieurs choix de réponses : Bunton et al., 2007 ; Lansford & Liss., 2014 ; Lansford et al., 2016 ; Van der Graaff et al., 2009 ; Zyski & Weisiger, 1987).

De fait, au vu des taux de précision obtenus par les juges à l'expérience de classification auditivo-perceptive globale des MSDs, ce type de raisonnement pourrait contribuer à faciliter le diagnostic différentiel des MSDs. Les cliniciens seraient alors ainsi guidés étape par étape à partir d'un arbre décisionnel. La figure 5 en constitue une proposition. Elle a été conceptualisée à partir de nos résultats et des données de la littérature malgré les manques théoriques évoqués en introduction générale.

Ainsi, face à un trouble de la communication verbale, avant de s'engager sur la piste d'un diagnostic de MSD, le clinicien devra se demander s'il est bien face à un trouble de nature neurologique et exclure les signes relatifs à la variabilité inter- et intra-individuelle de la parole neurotypique, ceux liés à un trouble organique (par exemple, fente palatine), ou à un déficit neurodéveloppemental persistant ou potentiellement décompensé (tels les troubles de la fluence de la parole neurodéveloppementaux : bégaiement, bredouillement, ou les troubles du développement des sons de la parole), ou encore à un trouble neurofonctionnel concernant la voix et la parole. Il lui faudra également faire la part des manifestations de déficits associés (langagiers, cognitifs, perceptifs, sensoriels…) et des atteintes neuropsychiatriques (troubles comportementaux et émotionnels) se répercutant sur les différentes dimensions de parole.

Une fois déterminée la nature neurologique du trouble sur la base des données médicales, biologiques, d'imageries ou d'explorations neurophysiologiques, de l'examen neurologique clinique et plus largement sensitivo-moteur, le clinicien interrogera le(s) niveau(x) de production langagière et/ou de production motrice de la parole concernant le(s) trouble(s) verbaux communicationnel(s) d'étiologie neurologique. Si le trouble est lié à un déficit de la production motrice de la parole et non à un déficit des niveaux

langagiers (aphasie) ou encore de l'incitation verbale (mutisme akinétique), le diagnostic pourra s'orienter vers un MSD.

Ensuite, une fois le MSD repéré sur la base d'une atteinte des dimensions de voix, de résonance ou de respiration pour la phonation et/ou sur la présence de distorsions ou d'allongements de sons, il sera question d'en définir le type : AoS ou dysarthrie.

Enfin, le sous-type sera précisé selon l'étiologie pour les AoS (non neurodégénérative ou neurodégénérative : prosodique, phonétique ou mixte) (Duffy et al., 2020 ; Utianski et al., 2018) et selon le système neurophysiologique atteint pour les dysarthries (Auzou, 2007 ; Darley et al., 1975 ; Duffy, 2019). Pour les distinguer, le clinicien pourra s'orienter vers un diagnostic d'AoS au vu de la présence d'un ralentissement du débit de parole en rapport avec des allongements de sons et/ou des intervalles intersyllabiques (D'Alessandro et al., 2019) et/ou de l'existence d'une dissociation entre un débit articulatoire plus rapide en DDK AMRs qu'en DDK SMRs (Lancheros et al., 2022) et/ou de la survenue de distorsions de sons en lien avec une réduction de la coarticulation (D'Alessandro et al., 2019).

Le raisonnement pourra pour certains locuteurs s'avérer moins aisé du fait de l'association fréquente avec d'autres troubles verbaux communicationnels.

L'ajout d'un entraînement spécifique, à la fois perceptif global et phonético-acoustique discret, permettrait d'améliorer le diagnostic et d'affiner la caractérisation des sous-types de MSDs au vu de la complémentarité de ces approches et de l'effet du degré d'expertise des auditeurs de l'étude 1. Il pourra être réalisé sur des échantillons de parole pathologique mais aussi sur ceux de locuteurs neurotypiques compte tenu de la variabilité intra- et interindividuelle de leur parole. Une réflexion, d'ordre didactique, portant sur les méthodes d'apprentissage (choix des échantillons : locuteurs, tâches de parole, méthodes et supports d'apprentissage perceptif, mode d'exposition) devrait alors être amorcée (pour exemple, concernant l'entraînement perceptif global : Borrie et al., 2012).

Parmi les développements ultérieurs de MonPaGe-2.0.s, l'intégration d'autres marqueurs de parole et/ou d'autres modules dans MonPaGe-2.0.s pourrait être envisagée en exploitant les différentes propriétés des tâches de parole contrôlées au moment de la conception de MonPaGe-2.0.s, ainsi que d'autres dimensions de parole (par exemple, la résonance nasale). L'outil pourrait viser dans une future version une plus grande sensibilité, notamment pour les formes très légères ; elles pourraient être détectées, entre autres, au moyen d'un paradigme de double tâche prenant en compte les propriétés des tâches de parole et non verbales qui influencent ses effets tels que retrouvés dans l'étude 3 chez les locuteurs neurotypiques. De fait, à l'heure actuelle, les tâches de parole de l'outil MonPaGe-2.0.s ne reproduisent pas les conditions - dont la demande cognitive - exercées sur la parole en vie quotidienne. Or, le locuteur dysarthrique ou présentant une apraxie de la parole se plaint généralement de difficultés éprouvées en situation écologique. Au moment où il parle, il se trouve fréquemment dans un environnement où il est simultanément confronté à d'autres stimuli. Il est aussi souvent en train de réaliser une autre action comme marcher, ce qui le place en situation de double tâche où le degré de demande attentionnelle se trouve modifié. Son déficit moteur de parole pourrait alors ne pas transparaître dans les tests effectués classiquement en condition isolée lors des évaluations, d'autant plus si son MSD est de sévérité très légère ou légère. Nous pourrions ainsi supposer que la plainte des locuteurs avec HD très légère non détectée par les auditeurs de l'étude perceptive (étude 1), ni par l'outil MonPaGe-2.0.s (étude 2), aurait pu potentiellement se manifester en condition de double tâche.

Enfin, à terme, les profils des marqueurs de MonPaGe-2.0.s retrouvés lors de l'évaluation des locuteurs avec MSD pourraient contribuer à la pose du diagnostic différentiel entre l'AoS et les différents sous-types de dysarthries.

À l'avenir, les systèmes de reconnaissance automatique de la parole pourraient également être des outils prometteurs en recherche et en clinique pour contribuer au diagnostic des locuteurs présentant un MSD et à l'évaluation de leur parole. Ils permettraient l'analyse de plus larges corpus et la prise en compte de signes survenant de manière intermittente en parole continue, avec la fatigue par exemple. En détectant d'autres marqueurs diagnostiques, ils pourraient faciliter la distinction entre AoS et dysarthries ou encore repérer des modifications légères lors des évaluations de suivi (Kodrasi et al., 2021). Les analyses issues de ces outils pourraient probablement être aussi complétées par des

données relatives aux biomarqueurs génétiques, sanguins, issus de la neuroimagerie (Josephs et al., 2021), ou par des données instrumentales (ultrasons, électropalatographie, articulographie électromagnétique, etc.). Des algorithmes combinant l'ensemble des marqueurs retenus, de diverses natures, viendraient alors faciliter et préciser le diagnostic et/ou permettre le suivi de l'évolution des troubles.

DECLARATION D'INTÉRETS

L'autrice a déclaré n'avoir aucun lien d'intérêt en relation avec cet article.

RÉFÉRENCES

Adams, S. G., Winnell, J., & Jog, M. (2010). Effects of interlocutor distance, multi-talker background noise, and a concurrent manual task on speech intensity in Parkinson's disease. *Journal of Medical Speech-Language Pathology, 18*(4), 1-9.

Allison, K. M., Cordella, C., Iuzzini-Seigel, J., & Green, J. R. (2020). Differential diagnosis of apraxia of speech in children and adults: A scoping review. *Journal of Speech, Language, and Hearing Research, 63*(9), 2952-2994. https://doi.org/10.1044/2020_JSLHR-20-00061

Audacity Team (2015). *Audacity® free audio editor and recorder* (version 2.1.0) [logiciel informatique]. https://audacityteam.org/

Auzou, P. (2007). L'évaluation clinique des dysarthries. Dans P. Auzou, V. Rolland-Monnoury, S. Pinto, & C., Özsancak (dir.), *Les dysarthries* (p. 189-195). Solal

Auzou, P., & Rolland-Monnoury, V. (2019). *BECD : batterie d'évaluation clinique de la dysarthrie* (2e éd.). Ortho Edition.

Bailey, D. J., & Dromey, C. (2015). Bidirectional interference between speech and nonspeech tasks in younger, middle-aged, and older adults. *Journal of Speech, Language, and Hearing Research, 58*(6), 1637-1653. https://doi.org/10.1044/2015_JSLHR-S-14-0083

Ballard, K. J., Azizi, L., Duffy, J. R., McNeil, M. R., Halaki, M., O'Dwyer, N., Layfield, C., Scholl, D. I., Vogel, A.P., & Robin, D. A. (2016). A predictive model for diagnosing stroke-related apraxia of speech. *Neuropsychologia, 81*, 129-139. https://doi.org/10.1016/j.neuropsychologia.2015.12.010

Boersma, P. & Weenink, D. (2021). *Praat: Doing phonetics by computer* (version 6.1.51) [logiciel informatique]. http://www.praat.org/

Borrie, S. A., McAuliffe, M. J., & Liss, J. M. (2012). Perceptual learning of dysarthric speech: A review of experimental studies. *Journal of Speech, Language and Hearing Research, 55*(1), 290-305. https://doi.org/10.1044/1092-4388(2011/10-0349)

Bunton, K., Kent, R. D., Duffy, J. R., Rosenbek, J. C., & Kent, J. F. (2007). Listener agreement for auditory-perceptual ratings of dysarthria. *Journal of Speech, Language, and Hearing Research,50*(6), 1481-1495. https://doi.org/10.1044/1092-4388(2007/102)

Chicherio, C., Genoud-Prachex, T., Assal, F., & Laganaro, M. (2019). *E-GeBAS: Electronic 534 Geneva Bedside Aphasia Scale* [logiciel informatique]. https://www.unige.ch/fapse/logotools/fr/adultes/e-gebas-electronic-geneva-bedside-aphasia-scala/

Code, C. (1998). Major review: Models, theories and heuristics in apraxia of speech. *Clinical Linguistics and Phonetics, 12*(1), 47–65. https://doi.org/10.3109/02699209808985212

D'Alessandro, D., Pernon, M., Fougeron, C., Laganaro, M. (2019, June 17-19). *Anticipatory VtoV coarticulation in French in several Motor Speech Disorders* [communication orale]. Third Phonetics and Phonology in Europe (PAPE), Lecce. https://hal.science/hal-02427864v1

Darley, F. L., Aronson, A. E., & Brown, J. R. (1975). *Motor speech disorders.* WB Saunders Company.

Dick, A. S., Garic, D., Graziano, P., & Tremblay, P. (2019). The frontal aslant tract (FAT) and its role in speech, language and executive function. *Cortex, 111*, 148-163. https://doi.org/10.1016/j.cortex.2018.10.015

Dromey, C., & Bates, E. (2005). Speech interactions with linguistic, cognitive, and visuomotor tasks. *Journal of Speech, Language, and Hearing Research, 48*(2), 295-305. https://doi.org/10.1044/1092-4388(2005/020)

Duffy, J. R. (2019). *Motor speech disorders: Substrates, differential diagnosis, and management* (4e ed.). Elsevier Health Sciences.

Duffy, J. R., Utianski, R. L., & Josephs, K. A. (2020). Primary progressive apraxia of speech: From recognition to diagnosis and care. *Aphasiology, 35*(4), 560-591. https://doi.org/10.1080/02687038.2020.1787732

Duffy, J. R., Martin, P. R., Clark, H. M., Utianski, R. L., Strand, E. A., Whitwell, J. L., & Josephs, K. A. (2023). The apraxia of speech rating scale: Reliability, validity, and utility. *American Journal of Speech-Language Pathology, 32*(2), 469-491.https://doi.org/10.1044/2022_AJSLP-22-00148

Ebersbach, G., Dimitrijevic, M. R., & Poewe, W. (1995). Influence of concurrent tasks on gait: A dual-task approach. *Perceptual and Motor Skills, 81*(1),107-113. https://doi.org/10.2466/pms.1995.81.1.107

Eichhorn, J. T., Kent, R. D., Austin, D., & Vorperian, H. K. (2018). Effects of aging on vocal fundamental frequency and vowel formants in men and women. *Journal of Voice, 32*(5), 644.e1-644.e9. https://doi.org/10.1016/j.jvoice.2017.08.003

Fonville, S., van der Worp, H. B., Maat, P., Aldenhoven, M., Algra, A., & van Gijn, J. (2008). Accuracy and inter-observer variation in the classification of dysarthria from speech recordings. *Journal of Neurology, 255*(10), 1545-1548. https://doi.org/10.1007/s00415-008-0978-4

Fougeron, C., Delvaux, V., Ménard, L., & Laganaro, M. (2018). The MonPaGe_HA database for the documentation of spoken French throughout adulthood. Dans *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)* (p. 4301-4306). https://aclanthology.org/L18-1679/

Fournet, M., Pernon, M., Catalano Chiuvé, S., Lopez, U., & Laganaro, M. (2021). Attention in post-lexical processes of utterance production: Dual-task cost in younger and older adults. *Quarterly Journal of Experimental Psychology, 174*(11), 1852–1872. https://doi.org/10.1177/17470218211034130

Gorno-Tempini, M. L., Hillis, A. E., Weintraub, S., Kertesz, A., Mendez, M., Cappa, S. F., Ogar, M., Rohrer, J. D., Black, S., Boeve, B. F., Manes, F., Dronkers, N. F., Vandenberghe R., Rascovsky, K., Patterson, K., Miller, B. L., Knopman, D. S., Hodges, J. R., Mesulam, M. M., & Grossman, M. (2011). Classification of primary progressive aphasia and its variants. *Neurology, 76*(11), 1006-1014. https://doi.org/10.1212/WNL.0b013e31821103e6

Harmon, T. G., Jacks, A., & Haley, K. L. (2019). Speech fluency in acquired apraxia of speech during narrative discourse: Group comparisons and dual-task effects. *American Journal of Speech-Language Pathology, 28*(2S), 905-914. https://doi.org/10.1044/2018_AJSLP-MSC18-18-0107

Hickok, G. (2014). The architecture of speech production and the role of the phoneme in speech processing. *Language, Cognition and Neuroscience, 29*(1), 2–20. https://doi.org/10.1080/01690965.2013.834370

Ho, A. K., Iansek, R., & Bradshaw, J. L. (2002). The effect of a concurrent task on Parkinsonian speech. *Journal of Clinical and Experimental Neuropsychology, 24*(1), 36-47.https://doi.org/10.1076/jcen.24.1.36.972

Hybbinette, H., Östberg, P., & Schalling, E. (2021). Intra- and interjudge reliability of the apraxia of speech rating scale in early stroke patients. *Journal of Communication Disorders, 89*, 106076. https://doi.org/10.1016/j.jcomdis.2020.106076

Jonkers, R., Feiken, J., & Stuive, I. (2017). Diagnosing apraxia of speech on the basis of eight distinctive signs. *Canadian Journal of Speech Language Pathology and Audiology, 41*(3), 303-319. https://cjslpa.ca/files/2017_CJSLPA_Vol_41/No_03/CJSLPA_Vol_41_No_3_2017_Jonkers_et_al_303_319.pdf

Josephs, K. A., Duffy, J. R., Clark, H. M., Utianski, R. L., Strand, E. A., Machulda, M. M., Botha, H., Martin, P. R., Thu Pham, N. T, Stierwalt, J., Ali, F., Buciuc, M., Baker, M., Fernandez de Castro, C. H., Spychalla, A. J., Schwarz, C. G., Reid, R. I., Senjem, M. L., Jack Jr., C. R., …, & Whitwell, J. L. (2021). A molecular pathology, neurobiology, biochemical, genetic and neuroimaging study of progressive apraxia of speech. *Nature Communications, 12*, 3452. https://doi.org/10.1038/s41467-021-23687-8

Kahneman, D. (1973). *Attention and effort* (vol. 1063). Prentice-Hall.

Kalafat, M., Hugonot-Diener, L., & Poitrenaud, J. (2003). Standardisation et étalonnage français du "Mini Mental State"(MMS) version GRECO. *Revue de Neuropsychologie, 13*(2), 209-236.

Kemper, S., Schmalzried, R., Herman, R., Leedahl, S., & Mohankumar, D. (2009). The effects of aging and dual task demands on language production. *Aging, Neuropsychology, and Cognition, 16*(3), 241-259. https://doi.org/10.1080/13825580802438868

Kemper, S., Hoffman, L., Schmalzried, R., Herman, R., & Kieweg, D. (2011). Tracking talking: Dual task costs of planning and producing speech for young versus older adults. *Aging, Neuropsychology, and Cognition, 18*(3), 257-279.https://doi.org/10.1080/13825585.2010.527317

Kent, R. D. (1996). Hearing and believing: Some limits to the auditory-perceptual assessment of speech and voice disorders. *American Journal of Speech-Language Pathology, 5*(3), 7-23. https://doi.org/10.1044/1058-0360.0503.07

Kim, Y., Kent, R. D., & Weismer, G. (2011). An acoustic study of the relationships among neurologic disease, dysarthria type, and severity of dysarthria. *Journal of Speech, Language, and Hearing Research, 54*(2), 417-429. https://doi.org/10.1044/1092-4388(2010/10-0020)

Kodrasi, I., Pernon, M., Laganaro, M., & Bourlard, H. (2021, June). Automatic and perceptual discrimination between dysarthria, apraxia of speech, and neurotypical speech [communication par affiche]. Dans *ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (p. 7308-7312). IEEE. https://doi.org/10.1109/ICASSP39728.2021.9414283

Laganaro, M., Fougeron, C., Pernon, M., Levêque, N., Borel, S., Fournet, M., Catalano Chiuvé, S., 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

Lancheros, M., Pernon, M., & Laganaro, M. (2022). Is there a continuum between speech and other oromotor tasks? Evidence from motor speech disorders. *Aphasiology, 37*(5), 715-734.https://doi.org/10.1080/02687038.2022.2038367

Lansford, K. L., & Liss, J. M. (2014). Vowel acoustics in dysarthria: Speech disorder diagnosis and classification. *Journal of Speech, Language, and Hearing Research, 57*(1), 57-67. https://doi.org/10.1044/1092-4388(2013/12-0262)

Lansford, K. L., Berisha, V., & Utianski, R. L. (2016). Modeling listener perception of speaker similarity in dysarthria. *The Journal of the Acoustical Society of America, 139*(6), EL209-EL215. https://doi.org/10.1121/1.4954384

McNeil, M. R., Robin, D. A., & Schmidt, R. A. (2009). Clinical *Management of Sensorimotor Speech Disorders* (2e ed.). Thieme Medical Publishers.

Melle, N., & Gallego, C. (2012). Differential diagnosis between apraxia and dysarthria based on acoustic analysis. *The Spanish Journal of Psychology, 15*(2), 495-504.http://dx.doi.org/10.5209/rev_SJOP.2012.v15.n2.38860

Menin-Sicard, A., & Sicard, E. (2020). Outils de mesures et d'évaluation objective de la phonologie, de la fluence et de la prosodie avec DIADOLAB3. *Rééducation Orthophonique*, 281, 131-168. https://www.orthoedition.com/revues/n-281-le-bilan-orthophonique-4283

Miller, H. E., & Guenther, F. H. (2021). Modelling speech motor programming and apraxia of speech in the DIVA/GODIVA neurocomputational framework. *Aphasiology, 35*(4), 424-441. https://doi.org/10.1080/02687038.2020.1765307

Molloy, J., & Jagoe, C. (2019). Use of diverse diagnostic criteria for acquired apraxia of speech: A scoping review. *International Journal of Language & Communication Disorders, 54*(6), 875-893. https://doi.org/10.1111/1460-6984.12494

MonPaGe (2021). [logiciel informatique]. https://lpp.cnrs.fr/monpage/

Mumby, K., Bowen, A., & Hesketh, A. (2007). Apraxia of speech: How reliable are speech and language therapists' diagnoses? *Clinical Rehabilitation, 21*(8), 760-767. https://doi.org/10.1177/0269215507077285

Pashler, H., Johnston, J. C., & Ruthruff, E. (2001). Attention and performance. *Annual Review of Psychology, 52*(1), 629-651. https://doi.org/10.1146/annurev.psych.52.1.629

Pernon, M., Trocello, J.-M., Vaissière, J., Cousin, C., Chevaillier, G., Rémy, P., Kidri-Osmani, K., Fougeron, C., & Woimant, F. (2013). Le débit de parole du patient wilsonien dysarthrique peut-il être amélioré en condition de double tâche? *Revue Neurologique, 169*(6-7), 502-509. https://doi.org/10.1016/j.neurol.2012.12.003

Pernon, M., Fougeron, C., Beurton, C., Bodiguel, L., Poujois, A., & Woimant, F. (2017, 29-30 juin). *Mesures de performance en production, de l'effet de la tâche et de l'attention sur les paramètres temporels de la parole de locuteur.rices contrôles et de locuteur.rices dysarthriques wilsoniens* [communication orale]. 7e Journées de Phonétique Clinique, Paris.

Pernon, M. (2018, 10 novembre). *Étude de la parole en double tâche : vers l'intégration de l'attention et des fonctions exécutives dans la prise en charge logopédique des troubles moteurs acquis de la parole ?* [communication orale]. De la recherche à la clinique logopédique : Apports des recherches menées au sein de la Maîtrise Universitaire en Logopédie, Genève.

Pernon, M., Fournet, M., Fougeron, C., & Laganaro, M. (2019, August 5-9). *Dual-task effects on speech and non-verbal tasks according to task properties* [communication par affiche]. Proceedings of the 19th International Congress of Phonetic Sciences, Melbourne, Australia. https://hal.science/hal-02427703v1

Pernon, M., Lévêque, 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, 169-198.

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.https://doi.org/10.1044/2022_JSLHR-21-00519

Pernon, M. (2022a). Parole - Sémiologie et évaluation clinique des troubles moteurs acquis de la parole : apraxies de la parole et dysarthries. Dans C. Sainson, C. Bolloré, & J. Trauchessec (dir.), *Neurologie et Orthophonie, Tome 1 : Théories et évaluation des troubles acquis de l'adulte* (p. 277-304). De Boeck Supérieur.

Pernon, M. (2022b). *Apports des approches d'évaluation perceptive, acoustique et cognitive des locuteurs présentant une apraxie de la parole et/ou une dysarthrie.* [Thèse de doctorat, Université Sorbonne-Nouvelle]. HAL. https://theses.hal.science/tel-04573087v1

Pommée, T., Balaguer, M., Mauclair, J., Pinquier, J., & Woisard, V. (2021). Assessment of adult speech disorders: Current situation and needs in French-speaking clinical practice. *Logopedics Phoniatrics Vocology, 47*(2), 92-108. https://doi.org/10.1080/14015439.2020.1870245

Qualtrics® (2019). [logiciel informatique]. https://www.qualtrics.com

R Core Team (2019). *The R Project for statistical computing*. R Foundation. https://www.R-project.org/

Sicard, E., & Menin-Sicard, A. (2020a). *DIADOLAB, logiciel professionnel d'évaluation et de réhabilitation de la parole* (version 3.1) [logiciel informatique]. http://www.diadolab.org

Sicard, E., & Menin-Sicard, A. (2020b). *VOCALAB, logiciel professionnel d'évaluation et de réhabilitation de la phonation* (version 4.1) [logiciel informatique]. http://www.vocalab.org

Southwood, M. H., & Dagenais, P. (2001). The role of attention in apraxic errors. *Clinical Linguistics & Phonetics, 15*(1-2), 113-116. https://doi.org/10.3109/02699200109167641

Strand, E. A., Duffy, J. R., Clark, H. M., & Josephs, K. (2014). The apraxia of speech rating scale: A tool for diagnosis and description of apraxia of speech. *Journal of Communication Disorders, 51*, 43-50. https://doi.org/10.1016/j.jcomdis.2014.06.008

Tourville, J. A., & Guenther, F. H. (2011). The DIVA model: A neural theory of speech acquisition and production. *Language and Cognitive Processes, 26*(7), 952-981. https://doi.org/10.1080/01690960903498424

Trouville, R., Delvaux, V., Fougeron, C., Laganaro, M. (2021). *Logiciel d'évaluation de la parole (version screening) MonPaGe-2.0.s* [logiciel informatique]. https://lpp.cnrs.fr/monpage/

Utianski, R., Duffy, J. R., Clark, H. M., Strand, E. A., Botha, H., Schwarz, C. G., Machulda, M. M., Senjem, M. L., Spychalla, A. J., Jack, C. R., Petersen, R. C., Lowe, V. J., Whitwell, J. L., & Josephs, K. A. (2018). Prosodic and phonetic subtypes of primary progressive apraxia of speech. *Brain and Language, 184*, 54–65. https://doi.org/10.1016/j.bandl.2018.06.004

Van der Graaff, M., Kuiper, T., Zwinderman, A., Van de Warrenburg, B., Poels, P., Offeringa, A., Van der Kooi, A., Speelman, H., & De Visser, M. (2009). Clinical identification of dysarthria types among neurologists, residents in neurology and speech therapists. *European Neurology, 61*(5), 295-300. https://doi.org/10.1159/000206855

Van der Merwe, A. (1997). A theoretical framework for the characterization of pathological speech sensorimotor control. Dans M. R. McNeil (dir.), *Clinical management of sensorimotor speech disorders* (p. 3–18). Thieme Medical Publishers.

Van der Merwe, A. (2021). New perspectives on speech motor planning and programming in the context of the four-level model and its implications for understanding the pathophysiology underlying apraxia of speech and other motor speech disorders. *Aphasiology, 35*(4), 397-423. https://doi.org/10.1080/02687038.2020.1765306

Wickens, C. D. (2008). Multiple resources and mental workload. *Human factors, 50*(3), 449-455. https://doi.org/10.1518/001872008X288394

Whitfield, J. A., & Goberman, A. M. (2017). Speech motor sequence learning: Effect of Parkinson disease and normal aging on dual-task performance. *Journal of Speech, Language, and Hearing Research, 60*(6S), 1752-1765. https://doi.org/10.1044/2017_JSLHR-S-16-0246

Whitfield, J. A., Kriegel, Z., Fullenkamp, A. M., & Mehta, D. D. (2019). Effects of concurrent manual task performance on connected speech acoustics in individuals with Parkinson disease. *Journal of Speech, Language, and Hearing Research, 62*(7), 2099-2117.https://doi.org/10.1044/2019_JSLHR-S-MSC18-18-0190

Zeplin, J., & Kent, R. D. (1996). Reliability of auditory-perceptual scaling of dysarthria. Dans D. A. Robin, K. M. Yorkston, & D. R. Beukelman (dir.), *Disorders of motor speech: Assessment, treatment, and clinical characterization* (p. 145-154). Paul H. Brookes Publishing Company.

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. https://doi.org/10.1006/brln.2001.2614

Ziegler, W. (2009). Modelling the architecture of phonetic plans: Evidence from apraxia of speech. *Language and Cognitive Processes, 24*(5), 631-661. https://doi.org/10.1080/01690960802327989

Zyski, B. J., & Weisiger, B. E. (1987). Identification of dysarthria types based on perceptual analysis. *Journal of Communication Disorders, 20*(5), 367-378. https://doi.org/10.1016/0021-9924(87)90025-6

Méthode d'analyse des pauses pour l'évaluation de troubles cognitifs subtils appliquée au Cancer-Related Cognitive Impairment.

Auteurs :
Amélie Richard[1,2]
Fabrice Hirsch[1]
Sophie Jacquin-Courtois[2,3]

Affiliations :
[1]Université Montpellier 3, Praxiling, CNRS UMR 5267, Montpellier, France.
[2] Université de Lyon, Centre de Recherche en Neurosciences de Lyon, CNRS UMR 5292, INSERM U1028, Lyon, France.
[3] Hôpital Henry Gabrielle, Hospices Civils de Lyon, Lyon, France.

Autrice de correspondance :
Amélie Richard
amelie.richard@inserm.fr

Dates :
Soumission : 24/07/2024
Acceptation : 25/11/2024
Publication : 28/12/2024

Comment citer cet article :
Richard, A., Hirsch, F. & Jacquin-Courtois, S. (2024).Méthode d'analyse des pauses pour l'évaluation de troubles cognitifs subtils appliquée au Cancer-Related Cognitive Impairment. *Glossa*, 142, 49-64. https://doi.org/10.61989/qp2vb532

e-ISSN : 2117-7155

Licence :
© Copyright Amélie Richard, Fabrice Hirsch et Sophie Jacquin-Courtois, 2024
Ce travail est disponible sous licence Creative Commons Attribution 4.0 International.

Contexte. L'analyse de la parole permet de repérer des troubles cognitifs subtils, notamment car la prosodie contient des éléments fins, comme les pauses, qui peuvent être des marqueurs comportementaux de Troubles NeuroCognitifs. Cependant, l'absence de méthodes simples et détaillées compromet la faisabilité d'une telle analyse en clinique. Les Troubles NeuroCognitifs mineurs (TNCm) désignent un changement cognitif rapporté par un individu pour lequel l'autonomie est préservée. Cette appellation recouvre de nombreuses physiopathologies sous-jacentes, tel que le Cancer-Related Cognitive Impairment (CRCI). Ce trouble cognitif induit par un cancer et ses traitements, est un TNCm principalement caractérisé par une plainte mnésique (« j'oublie ce qu'on me dit ») et linguistique (« je cherche mes mots ») dont la fréquence et l'intensité ne sont pas corrélées aux scores des tests orthophoniques ou neuropsychologiques. Le manque d'outil d'évaluation suffisamment sensible pour déceler ce trouble subtil conduit souvent à un sous-diagnostic. De nouvelles méthodes, comme l'analyse de la parole, sont donc souhaitables pour évaluer le CRCI.

Objectifs. Les objectifs de cette étude sont (i) de proposer une méthode d'analyse des pauses applicable en clinique, (ii) d'identifier les personnes porteuses d'un CRCI grâce à cette méthode.

Méthode. Treize participantes post-cancer du sein et treize participantes contrôles saines ont été incluses. Toutes ont eu pour instruction de raconter une histoire à partir d'une séquence d'images qui leur était présentée. Leur production orale a été enregistrée, puis transcrite automatiquement avec Whisper et analysée sur les logiciels SPPAS et Praat. Les pauses silencieuses, pauses remplies (« euh ») et allongements vocaliques (« un petiiit garçon ») ont été annotés, puis ont fait l'objet d'une analyse statistique sur JASP.

Résultats. La durée des pauses silencieuses des participantes post-cancer du sein était significativement plus longue que celle des contrôles. En revanche, la durée des pauses remplies et des allongements vocaliques n'a pas montré de différence significative entre les deux groupes. De même, le taux de pauses (nombre d'occurrences / durée totale du discours) ne montre pas de différence significative entre les deux groupes.

Conclusions. Cette étude décrit en détails les étapes d'analyse des pauses en vue de son applicabilité clinique. Les résultats suggèrent que la durée des pauses silencieuses permet de différencier les participantes post-cancer ayant une plainte cognitive des participantes contrôles.

Mots-clefs : Troubles NeuroCognitifs mineurs ; Cancer-Related Cognitive Impairment ; pauses ; transcription automatique ; parole ; prosodie.

Method for the analysis of pauses for assessing subtle cognitive impairment applied to Cancer-Related Cognitive Impairment.

Background. Speech analysis can detect subtle cognitive impairment, particularly because prosody contains fine elements, such as pauses, which might be behavioral markers of NeuroCognitive Disorders. However, the absence of simple, detailed methods compromises the feasibility of such an analysis in clinics. Mild Neurocognitive Disorders (mNCD) refer to a cognitive change reported by an individual whose autonomy is preserved. This terminology covers many underlying pathophysiologies, such as Cancer-Related Cognitive Impairment (CRCI). CRCI is an mNCD characterized by a memory complaint (e.g., "I forget what I'm told") and a language complaint (e.g., "I search for my words") for which the frequency and intensity exceed the scores of speech-language tests or neuropsychological tests. The lack of sensitive tools to detect this subtle impairment often leads to its underdiagnosis. New methods such as speech analysis are needed to assess CRCI.

Objectives. The study aims are (i) to propose a method for analyzing pauses which is practical in clinical context, (ii) to identify participants with CRCI using this method.

Methods. Thirteen participants post breast cancer and thirteen healthy controls were included in the study. The participants were instructed to tell a picture-based story. Their narratives were recorded, then automatically transcribed with Whisper and analyzed using SPPAS and Praat software. Silent pauses, filled pauses (e.g., "uh"), and sustained vowels (e.g., "a smaall boy") were annotated, then examined on JASP for a statistical analysis.

Results. The duration of silent pauses in post-cancer participants was significantly longer than that of the controls. However, the duration of filled pauses and sustained vowels did not show a significant difference between the two groups. Similarly, the rate of pauses (i.e., number of occurrences/total speech duration) did not show a significant difference between the two groups.

Conclusions. This study provides a well-described method for pause analysis designed for the purpose of clinical context. The results suggest that the duration of silent pauses is a good marker for differentiating post-cancer participants from controls.

Key-words: Mild NeuroCognitive Disorders ; Cancer-Related Cognitive Impairment ; pauses ; automatic transcription ; speech ; prosody.

CONTEXTE

Cet article propose un protocole permettant de conduire une analyse de la prosodie chez des patients ayant des troubles cognitifs subtils. Cette méthode se veut une aide au diagnostic de troubles cognitifs subtils. Elle s'adresse aux orthophonistes diplômés, aux étudiants en orthophonie et aux chercheurs ou cliniciens s'intéressant aux questions relatives à la fluence.

Marqueurs linguistiques de troubles cognitifs

La littérature en aphasiologie, vasculaire ou neurodégénérative, a montré l'intérêt de l'analyse de la parole et de la voix des patients pour révéler les troubles cognitifs. En effet, son caractère non invasif, sa sensibilité aux changements cognitifs subtils et sa haute validité écologique en font un outil de choix pour évaluer la cognition des patients (Lanzi et al., 2023). La parole peut être étudiée à travers le prisme de la phonétique, c'est-à-dire des éléments segmentaux de la parole (les phonèmes), et celui de la prosodie (He et al., 2023), autrement dit des éléments supra-segmentaux de la parole (fréquence fondamentale (Hz), intensité (dB) et durée (ms)). La prosodie se révèle particulièrement pertinente pour investiguer les Troubles NeuroCognitifs mineurs (TNCm) car elle contient des éléments fins signalant une modification de la cognition d'un locuteur (Ivanova et al., 2023). Ces éléments prosodiques comprennent des paramètres liés à la qualité de la voix, tels que des variations d'amplitude intonative (Themistocleous et al., 2020), ainsi que des unités plus larges comme les pauses (Vincze et al., 2021). Parjane et al. (2021) ont comparé la production orale de discours descriptifs de patients ayant une Aphasie Progressive Primaire non-fluente agrammatique, à celle de patients ayant une Démence à Corps de Lewy et de participants contrôles sains. Les résultats montrent des pauses plus longues, plus fréquentes et une réduction de la fréquence fondamentale dans les deux groupes de patients par rapport aux locuteurs contrôles.

Une des problématiques soulevées par ces recherches est liée à la variabilité des méthodologies employées (type de discours, variables linguistiques, outils d'analyse), rendant les résultats entre études difficilement comparables. Dans un effort d'harmonisation, Angelopoulou et al. (2024) ont proposé une méthode d'analyse de la parole comparant deux types d'élicitation de discours descriptifs (une description d'images et une description autobiographique) chez une population saine. L'intérêt de leur étude réside principalement sur la description des mesures de dix-neuf variables de parole, bien que certains concepts comme « mot » ou « pause » ne soient pas explicités (les auteurs ne différencient pas les pauses dues à des contraintes physiologiques des pauses ayant une fonction démarcative ou d'hésitation).

Pour autant, la faisabilité de leur méthodologie en pratique clinique est questionnable. En effet, la méthode d'analyse proposée par cette équipe est entièrement manuelle. Or, les étapes de transcription et d'annotation sont chronophages et exigent une rigueur pour être réplicables, généralisables ou simplement interprétables, les rendant difficiles à appliquer en routine clinique. Une méthode détaillée et facile à exécuter dans l'exercice de la clinique serait donc souhaitable.

Analyse (semi-)automatique de la parole

Le choix des outils utilisés détermine la qualité d'une analyse de la prosodie. Cette section argumente en faveur de l'utilisation d'outils informatiques automatisés, hors-ligne et *open source*.

Importance des outils automatisés

L'analyse de la parole est une tâche longue et fastidieuse. Après avoir enregistré la production orale d'un patient, il est souvent utile de la transcrire soit de manière orthographique, soit de manière phonétique. La transcription manuelle est particulièrement chronophage. En effet, la transcription d'une minute de parole peut prendre jusqu'à douze minutes (Bryant et al., 2017). Cette étape est également coûteuse en termes de concentration puisqu'elle demande plusieurs écoutes de l'enregistrement afin de ne pas omettre de mots ou de disfluences produites par le locuteur. Par exemple, il est difficile de connaître le nombre exact de répétitions à la première écoute. Le processus d'écoutes multiples entraîne des biais liés à la fatigabilité auditive, des biais de subjectivité et d'habituation, augmentant ainsi les risques d'erreurs de transcription (comme un ajout de mots par effet d'amorçage).

L'utilisation d'outils dotés d'Intelligence Artificielle (outils informatisés qui automatisent une tâche cognitive) permet de réduire à la fois le temps consacré à la tâche de transcription et à ses

biais en apportant par exemple une première transcription « neutre ». Ces outils sont issus de technologies plus ou moins complexes, de la simple tâche automatisée (détection automatique des pauses silencieuses) aux modèles de réseaux neuronaux (transcription automatique) et apportent un gain de temps considérable pour les analyses de la parole. Pour autant, les outils d'IA peuvent montrer des performances variables, notamment selon la qualité audio de l'enregistrement. Il est donc indispensable d'effectuer une vérification manuelle afin de corriger les éventuelles erreurs commises par ces outils.

Importance des outils hors-ligne

L'enregistrement de la voix de patients ou de participants constitue une donnée identifiante qui doit être protégée à un haut niveau de sécurité (Commission Nationale de l'Informatique et des Libertés, 2016). C'est pourquoi le choix d'outils transparents en matière de gestion des données est crucial. Pour autant, les ressources disponibles sur internet ont parfois des politiques de confidentialité obscures, ce qui accroît le risque de fuite des données. Les outils hors-ligne décrits dans cette étude répondent aux critères de sécurité requis pour l'analyse de la parole au sens que les données sont traitées depuis leur lieu de stockage.

Importance des outils open source

Les logiciels *open source* sont des logiciels gratuits dont les codes de programmation sont accessibles à tout utilisateur. Ce type d'outil collaboratif non seulement garantit une totale transparence en matière de gestion des données, mais favorise également la reproductibilité scientifique et la fiabilité des résultats.

Troubles NeuroCognitifs mineurs

Les troubles neurocognitifs désignent une modification des performances cognitives constatée par l'individu ou ses proches, et sont dits mineurs lorsque l'autonomie de l'individu est préservée (American Psychiatric Association, 2022). Les TNCm couvrent un large éventail de physiopathologies sous-jacentes telles que les pathologies neurodégénératives (maladie d'Alzheimer à un stade précoce, Démence à Corps de Lewy, etc.), les pathologies vasculaires (comme les accidents vasculaires cérébraux) ou traumatiques (traumatisme crânien léger) et les contextes infectieux ou oncologiques (séquelles de Covid-19, VIH, troubles cognitifs liés au cancer, etc.). L'appellation TNCm recouvre également les syndromes comportementaux ou cognitifs tels que le Mild Cognitive Impairment, pour lesquels le diagnostic d'une étiologie n'est pas établi.

La prévalence des TNCm est difficile à estimer du fait de la diversité des possibles étiologies et physiopathologies sous-jacentes. Il est donc possible que ces chiffres soient sous-estimés. Dans le cadre du traumatisme crânien léger, 10 à 15% des patients peuvent développer un syndrome post-commotionnel persistant incluant des troubles cognitifs (Groupe de travail Traumatisme Crânio-Cérébral Léger -TCCL- en Ile-de-France, 2021). Dans une étude menée par Hommel et al. (2009) auprès de soixante-quatorze patients post-AVC mineur ayant un score au *Mini Mental State Examination* (MMSE, Folstein et al., 1975) dans la norme (> 27/30), 70% des patients (soit 52/74 patients) rapportaient des difficultés de fonctionnement social corrélées à un déficit en mémoire de travail. Ce résultat semble indiquer la présence élevée de possibles troubles cognitifs après un AVC mineur. Par ailleurs, les auteurs suggèrent que les tests de dépistage cognitif usuels manqueraient de sensibilité à repérer ces troubles fins.

Lors de consultations cliniques, les patients ou leur entourage peuvent rapporter l'apparition de difficultés cognitives discrètes mais invalidantes. Cependant, leur plainte est rarement corrélée aux scores des tests de repérage cognitif tels que le MMSE ou le *Montreal Cognitive Assessment* (MoCA, Nasreddine et al., 2005). De fait, les TNCm restent souvent sous-diagnostiqués et il est rappelé par la Haute Autorité de Santé que toute plainte cognitive inquiétante, en particulier chez le sujet âgé, doit faire l'objet d'une investigation approfondie (Haute Autorité de Santé, 2018). Néanmoins, cette même vigilance se doit d'être appliquée avec la même rigueur à une population possiblement plus jeune dans d'autres contextes (AVC mineur, traumatisme crânien léger, Cancer-Related Cognitive Impairment, etc.) dont l'enjeu sanitaire et médico-économique est important.

Cancer-Related Cognitive Impairment

Parmi les TNCm les plus fréquents, le Cancer-Related Cognitive Impairment (désormais CRCI) tient une place singulière qualitative et quantitative. En effet, il s'agit d'un trouble cognitif subtil induit

par un cancer (hors cancer du système nerveux central) et/ou ses traitements, qui peut concerner jusqu'à 75% des patients en oncologie (Janelsins et al., 2011). Outre son incidence élevée, il altère la qualité de vie des patients. Pour certains, la symptomatologie est durable, pouvant persister plus de deux ans après l'arrêt des traitements oncologiques. La permanence du trouble peut entraîner une limitation de la participation aux activités sociales et entraver la reprise professionnelle d'une population possiblement jeune. Le caractère subtil de l'expression de ses troubles en fait un handicap invisible souvent sous-diagnostiqué, alors même que des possibilités d'évaluation appropriée et de prises en charge existent.

Le CRCI est caractérisé par une plainte cognitive touchant plusieurs aspects de la cognition (mémoire, fonctions exécutives, attention, vitesse de traitement, langage) pouvant être objectivée par des questionnaires standardisés (comme le FACT-Cog de Wagner et al., 2009). Une des plaintes les plus fréquentes est le manque du mot. Malgré l'intensité de la plainte, les tests orthophoniques et neuropsychologiques actuels, généralement conçus pour évaluer des troubles plus sévères, peinent à révéler pleinement les difficultés rapportées par les patients. Des études comparant des patientes ayant reçu un traitement par chimiothérapie pour un cancer du sein et des participantes contrôles saines montrent une différence significative pour les compétences verbales (Amani et al., 2024 ; Jim et al., 2012). Toutefois, la taille d'effet de cette différence est relativement petite, ce qui signifie que les scores des patientes sont proches de ceux des contrôles. Dès lors, il est possible que les scores obtenus notamment aux épreuves de fluence verbale et de dénomination administrées en clinique, soient souvent dans les normes. De nouvelles approches et/ou de nouveaux outils sont donc à développer pour détecter ces troubles cognitifs subtils.

OBJECTIF ET HYPOTHÈSES

Concernant les marqueurs prosodiques du CRCI, les études pilotes menées par notre équipe ont montré l'intérêt de l'analyse des disfluences (Richard et al., 2022) et particulièrement des pauses dans la production discursive de patientes ayant eu un cancer du sein qui ont une plainte cognitive (Richard et al., 2023). L'analyse des pauses d'hésitation produites lors d'une tâche de narration élicitée par des images se révèle suffisamment sensible pour différencier des patientes post-cancer du sein ayant une plainte cognitive de participantes contrôles (Richard et al., 2024).

Les objectifs de la présente étude sont (i) de proposer une méthode d'analyse prosodique applicable en clinique, (ii) d'appliquer cette méthode à l'analyse des pauses chez des femmes ayant une plainte cognitive des suites d'un cancer du sein.

Nous émettons l'hypothèse que la méthode d'analyse que nous proposons permettrait de mettre en évidence un CRCI chez des patientes ayant eu un cancer du sein. Les pauses d'hésitation seraient plus fréquentes (en nombre/durée de discours) et plus longues (en ms) pour les participantes ayant une plainte cognitive liée au cancer que pour les participantes contrôles saines.

MÉTHODE

Démarches éthiques

La présente étude est issue de deux projets plus larges. Le premier a pour objectif d'évaluer l'efficacité d'un programme de réadaptation physique proposé à des femmes ayant eu un cancer du sein à l'issue de la phase principale de leurs traitements (ActiCog) et a été approuvé par le comité éthique des Hospices Civils de Lyon (n°22-5009). Le second projet porte sur le trouble cognitif induit par un cancer et ses traitements chez des femmes ayant eu un cancer du sein (DisCCo) et a été validé par le Comité de Protection des Personnes Est-III en août 2023. L'étude est enregistrée au numéro national 2023-A00267-38.

Population

Treize participantes ayant eu un cancer du sein (six participantes issues de l'étude ActiCog et sept participantes issues de l'étude DisCCo) et treize participantes contrôles saines (étude DisCCo) appariées en âge et en niveau de scolarisation ont été éligibles pour cette étude.

Critères d'inclusion

Les critères d'inclusion des participantes ayant eu un cancer (groupe Breast Cancer, désormais BC) étaient (i) d'avoir terminé les traitements curatifs (chirurgie, chimiothérapie, radiothérapie) depuis plus de quatre semaines et moins d'un an, hors hormonothérapie, (ii) d'avoir une vue et une

audition normales ou corrigées, (iii) d'avoir un score à la sous-échelle *Perceived Cognitive Impairment* (PCI) du questionnaire FACT-Cog inférieur ou égal à 55/72. Le questionnaire FACT-Cog (Wagner et al., 2009) est un questionnaire standardisé d'évaluation de la plainte cognitive des patients en oncologie pour lequel un seuil inférieur ou égal à 55/72 à la sous-échelle PCI indique la présence d'une plainte cognitive significative (Van Dyk et al., 2020).

Les critères d'inclusion des participantes contrôles (groupe Healthy Controls, désormais HC) étaient les mêmes que les participantes ayant eu un cancer du sein, excepté pour le score PCI du FACT-Cog (score > 55/72).

Critères d'exclusion

Les critères d'exclusion des participantes BC étaient (i) la présence de métastases, (ii) des antécédents de troubles du langage oral, (iii) des antécédents neurologiques ou psychiatriques, (iv) la prise de médicaments susceptibles d'entraîner une modification de la cognition.

En plus des critères d'exclusion énoncés ci-dessus, les participantes HC devaient ne jamais avoir eu de cancer.

Tâche de narration

Il était demandé aux participantes de produire une tâche de discours narratif à partir d'une séquence de cinq images (Bézy et al., 2016). Les images racontent l'histoire d'une famille qui arrive à la plage. Le père fait de la planche à voile pendant que son fils et sa femme sont sur le rivage. Soudain, le vent se lève et le père tombe à l'eau. La dernière image dépeint la mère et le fils alarmés, la mère appelant les secours au téléphone. Les participantes ont pour consigne de retracer l'histoire en images qu'elles voient, sans limite de temps.

Matériel et conditions d'enregistrement

Les productions discursives des participantes à l'étude ActiCog ont été enregistrées dans une pièce calme de l'hôpital de rééducation Henry Gabrielle (Hospices Civils de Lyon, Saint Genis Laval, France) avec un micro-cravate omnidirectionnel Rode Lavalier Go® relié à un ordinateur. Le micro était placé à quinze centimètres de la bouche des participantes.

Les productions discursives des participantes à l'étude DisCCo ont été enregistrées dans une cabine anéchoïque du Centre de Recherche en Neurosciences de Lyon (Bron, France) avec un micro-cravate omnidirectionnel Shure® relié à une carte son de marque M-Audio® et branché à un ordinateur. Le micro était placé à quinze centimètres de la bouche des participantes.

Paramètres linguistiques

Nous avons analysé les pauses silencieuses ainsi que les pauses remplies et allongements vocaliques liés à l'hésitation. Nous avons analysé les pauses égales ou supérieures à 200 ms. Ce seuil tient en effet compte de paramètres liés à la production de la parole, comme la durée des silences acoustiques des consonnes occlusives qui est inférieure à cette valeur. Il correspond également à une limite perceptive à partir de laquelle la pause est significativement perçue (Candea, 2000 ; Goldman et al., 2010). Toutes les pauses silencieuses au-dessus de ce seuil ont été incluses. Pour être considérés comme des pauses remplies et des allongements vocaliques, les phonèmes de plus de 200 ms devaient avoir une courbe intonative plate (Campione & Véronis, 2002 ; Duez, 2001). Suivant ces critères, nous avons identifié trois types de pauses :

- Les pauses silencieuses : segments caractérisés par l'absence de parole, de toux ou de bâillement d'une durée supérieure ou égale à 200 ms,

- Les pauses remplies : segments voisés de type « euh », « hum » d'une durée supérieure ou égale à 200 ms ayant une courbe intonative (F0) plate,

- Les allongements vocaliques : tenue d'un phonème vocalique d'une durée supérieure ou égale à 200 ms dont la courbe intonative (F0) est plate.

Pour chaque type de pauses, nous avons extrait la durée moyenne, médiane et totale, ainsi que la fréquence totale pour calculer le taux de pauses (nombre d'occurrence/durée totale du discours, pauses silencieuses incluses).

Analyse des pauses

La section suivante décrit les différentes étapes de l'analyse prosodique réalisée à partir du signal acoustique. Les différentes phases de la chaîne de traitement peuvent être conduites indépendamment pour adapter l'analyse aux besoins cliniques ou au focus d'une étude. Le temps de réalisation de cette procédure dépend largement de la durée et du nombre d'extraits

audio analysés, ainsi que du niveau de profondeur de l'analyse. Pour cette étude, l'analyse des vingt-six enregistrements représentant trente-cinq minutes de parole a duré environ deux heures dont quatre-vingts minutes de transcription automatique.

Nous avons rédigé un tutoriel complet détaillant chaque étape de notre analyse, disponible sur la plateforme OSF, fichier Protocole_annotation_pauses[1].

Les tutoriels que nous proposons se veulent accessibles à tous et ne requièrent pas de compétences particulières en informatique.

Pré-traitement du signal acoustique

Avant de commencer nos analyses prosodiques, nous avons segmenté l'enregistrement audio pour ne conserver que la production narrative en supprimant les parties qui ne seront pas étudiées (comme l'intervention de l'examinateur). Cette étape a été effectuée avec le logiciel libre Audacity (Audacity Team, 1999). Cette étape a duré environ cinq minutes.

Transcription

Nous avons réalisé les transcriptions automatiquement avec le modèle large V3 du code Whisper (Radford et al., 2022). Whisper est une commande *open source* dédiée à la reconnaissance automatique de la parole. Une de ses fonctions consiste à transcrire un discours oral en texte rédigé à partir du signal acoustique. Une description de l'installation et de l'utilisation de Whisper (testée avec deux volontaires naïfs de la programmation informatique) est disponible sur la plateforme OSF, fichier Installation_Whisper_SPPAS_Praat[2].

Nous avons vérifié chaque texte afin de corriger les erreurs de transcription. Les erreurs étaient peu nombreuses et concernaient surtout les noms propres et les troncatures de mots pour les enregistrements de qualité moyenne. Nous avons ensuite ajouté manuellement les pauses silencieuses (#) et les pauses remplies (euh). Il est important de noter que Whisper supprime par défaut toutes les disfluences (pauses, répétition, faux-départ, etc.).

Enfin, nous avons automatiquement nettoyé les données en supprimant par exemple les marques de ponctuation ou en remplaçant les chiffres par leur équivalent en lettres, grâce à un code Python développé pour cette étude. Ce code est disponible sur la plateforme OSF, fichier transcript_cleaner.py[3].

Cette étape a duré environ une heure et demie dont quatre-vingts minutes de transcription automatique.

Annotation des pauses

Nous avons annoté les pauses en utilisant le logiciel SPPAS (Bigi, 2015). SPPAS est un logiciel *open source* spécifiquement développé pour l'analyse de la parole. Il fournit de nombreuses analyses automatiques telles que la détection des pauses silencieuses, la conversion graphème-phonème ou encore la syllabation de la transcription alignée sur le signal acoustique. Son interface graphique et son caractère interopérable (possibilité de fonctionner avec d'autres logiciels) en font un outil autant applicable en recherche qu'en clinique. Pour affiner nos analyses, nous avons également utilisé le logiciel libre Praat (Boersma & Weenink, 2024). Celui-ci a été conçu pour l'analyse du signal acoustique et permet de visualiser avec précision le signal pour en extraire de nombreuses variables prosodiques (la fréquence fondamentale p. ex.).

Le paragraphe suivant décrit les étapes que nous avons réalisées pour la phase d'annotation du corpus. L'ensemble de ces étapes a duré environ trente minutes.

Nous avons d'abord chargé les enregistrements (format `.wav`) et les transcriptions (format `.txt`) dans SPPAS. Depuis l'onglet « Annotation », nous avons recherché les unités inter-pausales (*Inter Pausal Units*, IPUs), puis lancé la complétion automatique de la transcription. La première fonction, « Search for IPUs », détecte automatiquement les pauses silencieuses et la seconde, « Fill in IPUs », complète les espaces vides par la transcription des segments de parole. Ces fonctions sont accessibles via le bouton « STANDALONE ». Nous avons choisi le format de sortie « `.TextGrid` », format traité par Praat.

Nous avons ensuite procédé à une étape de vérification en ouvrant les fichiers `.TextGrid` dans Praat. Nous avons réajusté les frontières des pauses silencieuses quand celles-ci ne correspondaient pas au signal acoustique (cf. figure 1). Les erreurs d'alignement étaient nombreuses pour les enregistrements de qualité moyenne (avec bruit de fond) et rares pour les enregistrements de bonne qualité.

[1,2,3] https://osf.io/sqwnz/?view_only=8947f20e707f476c9cf44a35ca9d7c32

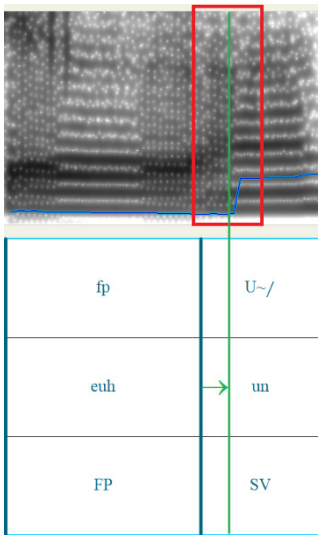


FIGURE 1 : Exemple d'erreur d'alignement. La pause « fp » (filled pause, pause remplie) est partiellement tronquée et la frontière (barre bleue) doit être tirée vers la droite jusqu'à la fin de l'émission du phonème (barre verte).

Dans l'onglet « Annotation » de SPPAS, nous avons lancé les fonctions « Text Normalization », « Phonetization » et « Alignment » simultanément. Cette étape permet de segmenter l'audio en phonèmes à partir du texte et d'aligner temporellement les phonèmes sur le signal acoustique. De cette manière, les allongements vocaliques et les pauses remplies sont facilement repérables par leur durée égale ou supérieure à 200 ms.

Nous avons effectué une nouvelle vérification dans Praat en ouvrant les fichiers `.TextGrid` nommés *-palign* pour nous assurer que les phonèmes étaient correctement alignés avec le signal acoustique.

Dans SPPAS, onglet « Analyze », nous avons ouvert les fichiers `.TextGrid` intitulés *-palign*, et nous avons sélectionné la ligne d'annotation (ou *tier*) « TokenAlign ». Nous avons appliqué un filtre (*single filter*) pour créer un nouveau *tier* nommé « pauses ». Ce *tier* contient les phonèmes vocaliques et les pauses `(U~/, A/, O/, @, i, E, e, j, H, a~, O~, y, 2, 9, u, fp, #` selon les conventions de SPPAS) d'une durée égale ou supérieure à 200 ms. Dans Praat, nous avons uniquement conservé les annotations des allongements vocaliques dont la courbe intonative est plate.

Enfin, nous avons harmonisé l'ensemble des annotations en remplaçant les annotations déjà existantes du *tier* « pauses » par les étiquettes suivantes : **SP** pour les pauses silencieuses (*Silent Pauses*), **FP** pour les pauses remplies (*Filled Pauses*) et **SV** pour les allongements vocaliques (*Sustained Vowels*). Pour cela, nous avons lancé le script de remplacement de texte développé par DiCanio (2011) dans Praat.

Accord inter-annotateur

L'accord inter-annotateur consiste à donner un certain pourcentage des données (environ 20%) ainsi que son protocole d'annotation (un descriptif des éléments à annoter) à une ou plusieurs personnes n'ayant pas de connaissances du corpus afin de vérifier la validité et la fiabilité du processus d'annotation. Il se calcule à l'aide du coefficient κ de Cohen, dont le score s'échelonne de 0 à 1. Un score égal ou supérieur à .81 atteste la validité et la fiabilité des annotations. En clinique, cette étape permet d'assurer une annotation comportant le moins de biais de subjectivité possible. En effet, la connaissance du patient et de ses troubles peut conduire à surestimer ou sous-estimer le nombre de pauses. Il est vrai cependant que l'accord inter-annotateur peut être difficile à réaliser en clinique (par manque de disponibilité de ses collègues p. ex.), aussi cette étape peut être optionnelle.

Nous avons confié 20% de nos données à deux annotatrices cliniciennes non-orthophonistes, naïves de l'analyse de la parole, pour comparer nos annotations des pauses. Nous leur avons fourni en amont un document détaillé de notre système d'annotation. Notre accord inter-annotateur a atteint 98,71% d'accord pour l'annotation des pauses silencieuses, soit un agrément presque parfait (κ = .97), 90,98% pour l'annotation des pauses remplies, soit un agrément presque

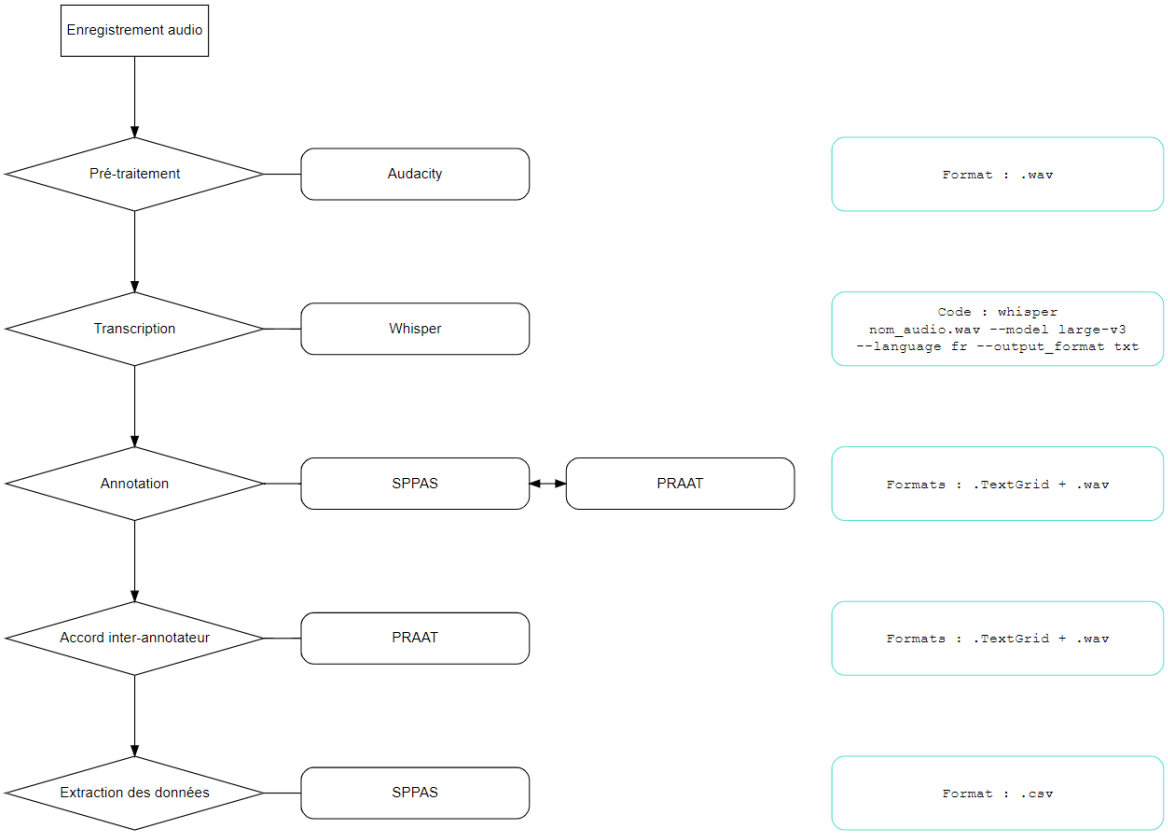


FIGURE 2 : Diagramme des étapes et des logiciels utilisés pour l'annotation de la parole.

parfait (κ = .81), et 86,88% pour l'annotation des allongements vocaliques, soit un agrément substantiel (κ = .74).

La figure 2 synthétise les différentes étapes que nous avons réalisées pour annoter la parole des participantes.

Extraction des données prosodiques

Nous avons extrait le nombre, la durée totale, moyenne, médiane et l'écart-type des trois types de pauses pour chaque échantillon audio à partir de SPPAS. Dans l'onglet « Analyze », nous avons sélectionné toutes les *tiers* « pauses », puis ouvert l'option « Statistics » pour accéder aux données sous forme de tableur. Cette étape a duré deux minutes.

Évaluation psychologique et cognitive

L'ensemble des participantes a bénéficié d'une évaluation psychologique et cognitive afin de contrôler les variables confondantes susceptibles d'influencer la production de pauses.

Elles ont également complété la version française du questionnaire FACT-Cog v3 (Costa et al., 2018 ; Joly et al., 2012) pour évaluer la présence d'une plainte cognitive significative (score inférieur ou égal à 55/72 à la sous-échelle *Perceived Cognitive Impairment*, PCI).

Les femmes ayant pris part à l'étude ont complété l'échelle *Hospital Anxiety Depression* (HAD, Zigmond & Snaith, 1983). Ce questionnaire comprend deux sous-échelles destinées à évaluer l'anxiété (HAD-A) et la dépression (HAD-D). Un score supérieur à 11/21 à l'une de ces sous-échelles atteste de la présence d'anxiété ou de dépression.

Elles ont ensuite complété le *Functional Assessment of Chronic Illness Therapy – Fatigue* (FACIT-F, Bonomi et al., 1996 ; Cella et al., 1993). Ce questionnaire est spécialement conçu pour évaluer la fatigue globale liée au cancer. Plus le score est élevé, moins la fatigue est importante. Il est à noter qu'il n'existe pas de seuil standardisé pour cette batterie de questions.

Enfin les participantes ont passé le test de repérage cognitif MoCA (Nasreddine et al., 2005). Ce test rapide à administrer est conçu pour détecter des troubles de la cognition en évaluant plusieurs

	id	site	group	age	complaint	anxiety	depression	cognition	FP_mean	SP_mean	SV_mean	FP_median
1	AL04	hospital	BC	41	38	0	2	26	0.464137214	1.259361109	0.464098824	0.25
2	AL08	hospital	BC	38	39	4	3	26	0.505	0.767647059	0.429166667	0.505
3	AL10	hospital	BC	38	31	11	4	30	0.568072969	0.756917516	0.318287043	0.48893414
4	AL13	hospital	BC	45	29	9	13	21	0.639166667	1.425927875	0.401111111	0.64
5	AL14	hospital	BC	39	49	6	11	29	0.626744848	0.799075957	0.393893653	0.5225

FIGURE 3 : Rangement des données de cinq participantes dans JASP.

domaines cognitifs (mémoire, capacités visuo-spatiales, langage, etc.). Un score supérieur ou égal à 26/30 est considéré comme normal.

Analyses statistiques

Les analyses statistiques de cette étude ont été réalisées avec le logiciel JASP (JASP Team, 2023). Ce logiciel présente l'avantage d'être gratuit, hors ligne, *open source* et simple d'utilisation. Son interface graphique rend accessibles plusieurs modèles de statistiques fréquentistes paramétriques et non paramétriques, ainsi que leurs équivalents Bayésiens (cf. https://jasp-stats.org/how-to-use-jasp/ pour une introduction à JASP). Nos données sont rangées dans un tableur au format .csv. Chaque colonne correspond à une variable et chaque ligne à une participante. La figure 3 montre le rangement des données après importation dans JASP.

L'effectif de notre étude est petit, raison pour laquelle nous avons choisi des tests non-paramétriques pour nos analyses statistiques. Nous avons effectué des tests U de Mann-Whitney (Mann & Whitney, 1947) pour comparer (1) l'âge, les scores au test de repérage cognitif et aux questionnaires des participantes BC par rapport aux participantes HC, ainsi que (2) les variables prosodiques des deux groupes. Enfin, nous avons effectué une corrélation de Spearman (Spearman, 1904) pour vérifier la possibilité d'un effet des variables psychologiques (anxiété, dépression, fatigue) sur les variables prosodiques.

L'ensemble des graphiques a été réalisé sur R studio (RStudio, BPC, 2020) avec les packages ggplot2 (Wickham et al., 2016) et GridExtra (Kassambara, 2016).

RÉSULTATS

Caractéristiques du corpus

Notre corpus est composé de 26 échantillons de parole pour une durée totale de 35 minutes (moyenne par échantillon = 82 s ± 41s). Au total, nous avons annoté 1222 pauses d'hésitation.

Données démographiques

Le tableau 1 montre les résultats du test U de Mann-Whitney réalisé pour la comparaison des variables démographiques et psychologiques entre les groupes. Comme attendu l'âge (p = .456) et le score MoCA ne révèlent pas de différence significative (p = .407), contrairement au score PCI du FACT-Cog (p < .001). Les participantes BC ont des scores significativement plus élevés que les contrôles aux questionnaires d'anxiété (p = .031) et de dépression (p = .009). Ces scores restent toutefois en-deçà des seuils pathologiques. Les participantes BC ont un score significativement

TABLEAU 1 : Comparaison des variables démographiques et des scores au test cognitif et aux questionnaires.

Variable	BC *moyenne (ET)*	HC *moyenne (ET)*	*p*-value
Âge	46 (9.9)	44 (10.6)	0.456
MoCA (/30)	27 (2.6)	27 (2.3)	0.407
FACT-Cog PCI (/72)	31 (10.7)	67 (4.8)	**< .001***
HAD-A (/24)	10 (5.0)	6 (4.2)	**0.031***
HAD-D (/24	7 (4.5)	3 (3.3)	**0.009***
FACIT-F (/52)	25 (8.6)	44 (7.1)	**< .001***

* = différence significative.

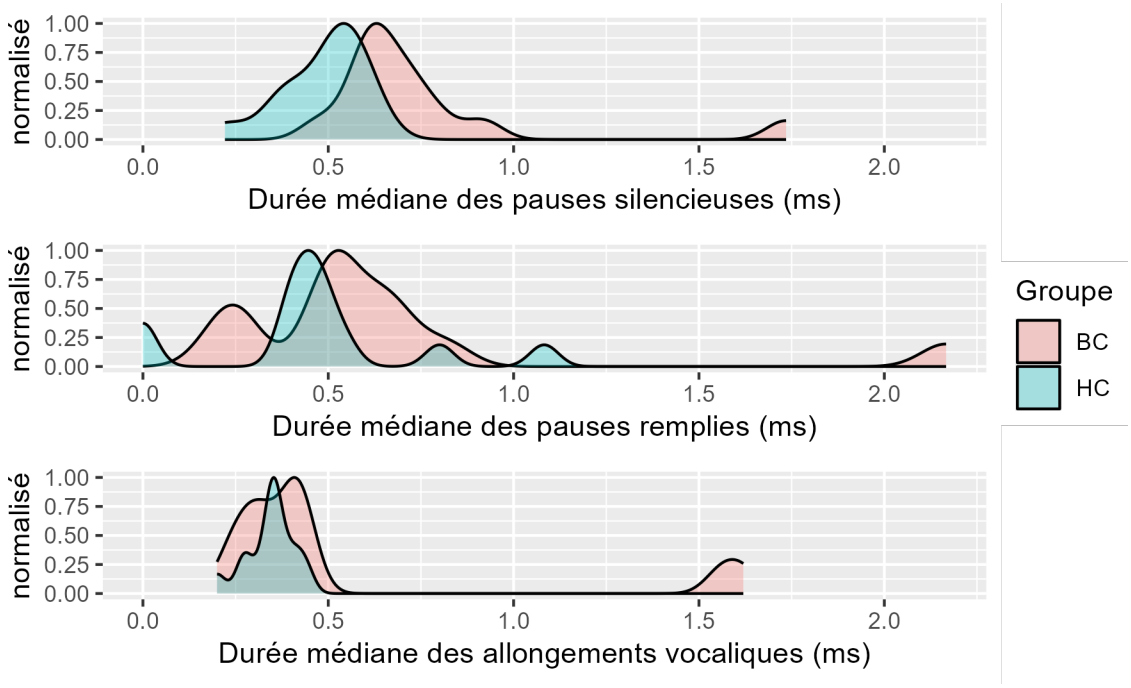


FIGURE 4 : Distribution de la durée médiane des pauses.

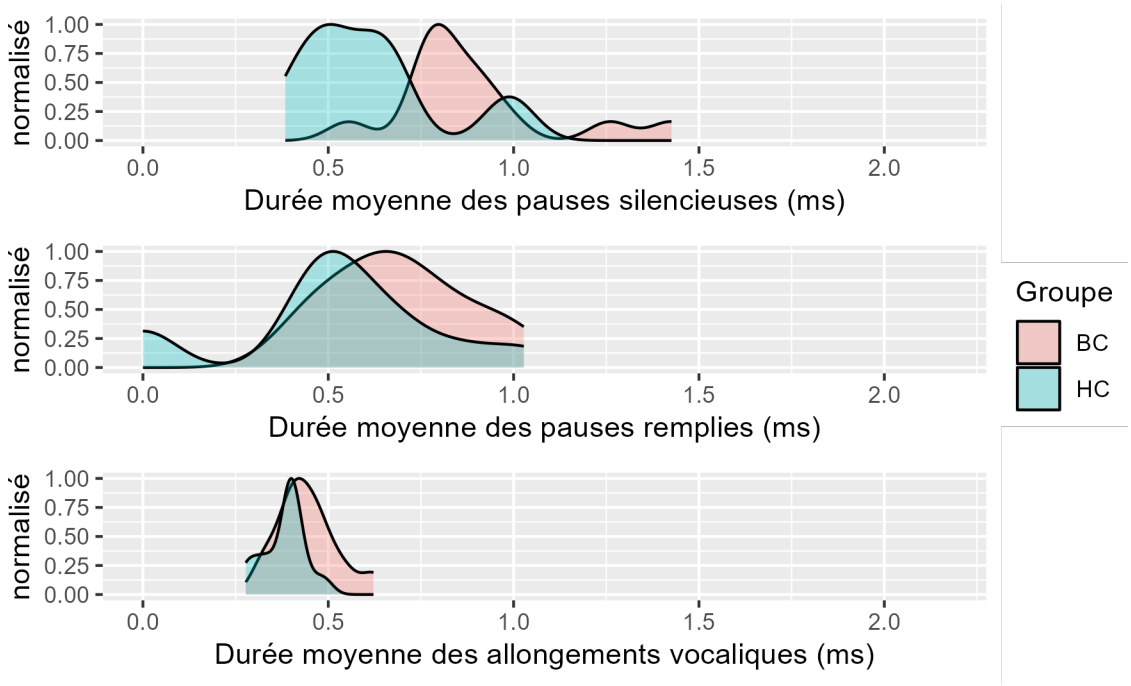


FIGURE 5 : Distribution de la durée moyenne des pauses.

plus faible au questionnaire de fatigue ($p < .001$), ce qui suggère la possible présence de fatigue chez cette population.

Résultats de l'analyse des pauses

La distribution des pauses est asymétrique avec une densité plus élevée chez les deux groupes pour les pauses de courte durée (~500 ms). La figure 4 montre la distribution de la durée médiane des pauses pour chaque type de pause.

La figure 5 montre la distribution de la durée moyenne des pauses pour chaque type de pause.

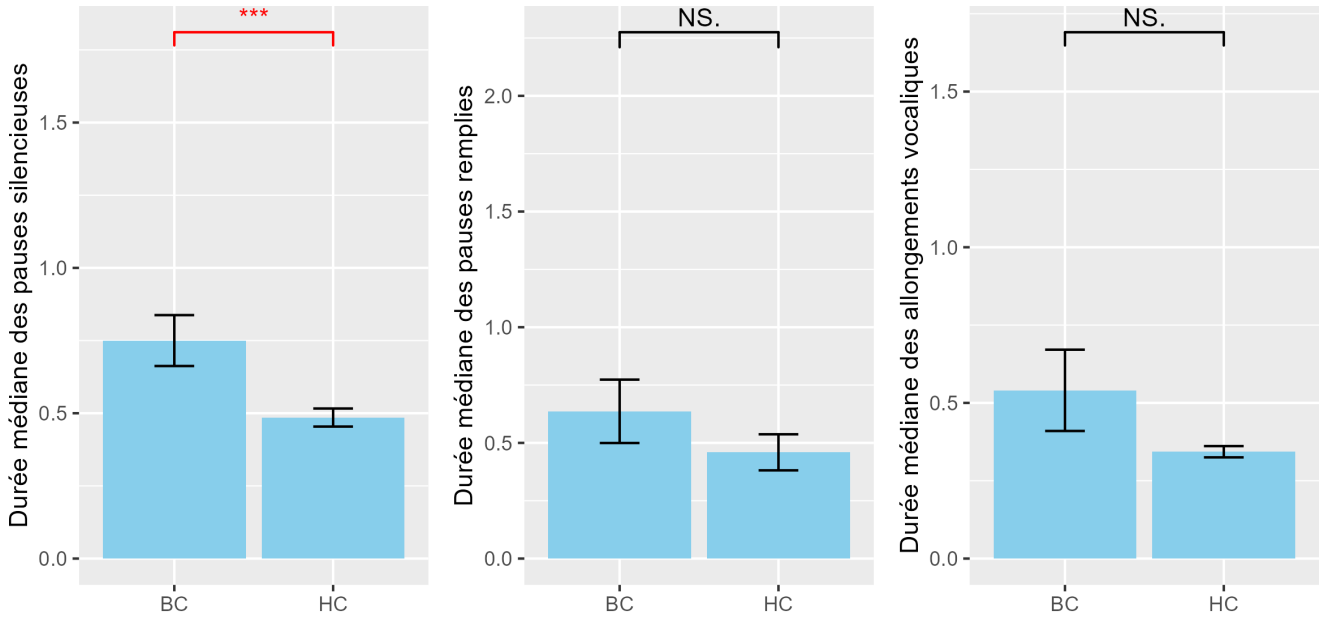


FIGURE 6 : Diagramme en barres de la durée médiane des pauses silencieuses, pauses remplies et allongements vocaliques pour les groupes Breast Cancer (BC) et Healthy Controls (HC). Les barres indiquent l'erreur-type. Les étoiles indiquent une différence significative entre les deux groupes.

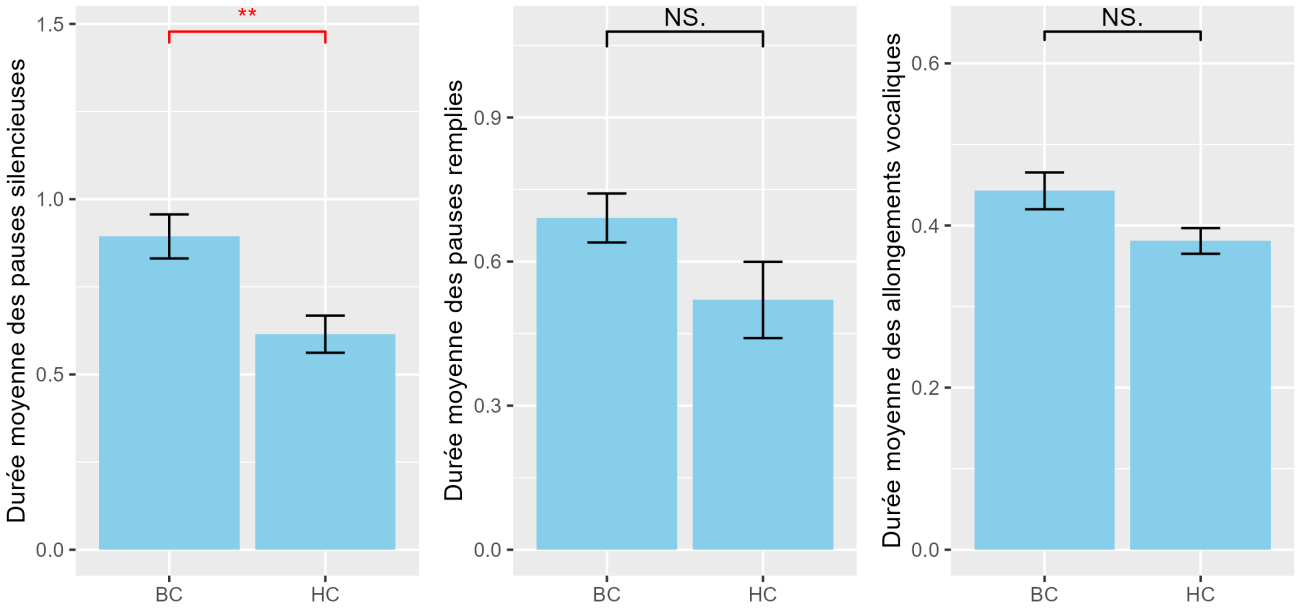


FIGURE 7 : Diagramme en barres de la durée moyenne des pauses silencieuses, pauses remplies et allongements vocaliques pour les groupes Breast Cancer (BC) et Healthy Controls (HC). Les barres indiquent l'erreur-type. Les étoiles indiquent une différence significative entre les deux groupes.

Durées médiane et moyenne des variables prosodiques

Nous avons réalisé un test U de Mann-Whitney pour comparer les durées médiane et moyenne des pauses silencieuses, des pauses remplies et des allongements vocaliques des participantes BC à celles des participantes HC. Les résultats indiquent des durées médiane et moyenne des pauses silencieuses significativement plus longues pour les participantes BC que pour les participantes HC (respectivement p = <.001 et p = .002). En revanche, les résultats ne montrent

pas de différence significative pour les durées médiane et moyenne des pauses remplies et des allongements vocaliques.

La figure 6 et la figure 7 montrent respectivement les durées médiane et moyenne des pauses silencieuses, pauses remplies et allongements vocaliques pour les groupes BC et HC.

Taux de pauses et d'allongements vocaliques

Les taux de pauses et d'allongements vocaliques par échantillon ont été calculés en divisant le nombre d'occurrences par la durée totale des échantillons.

Le tableau 2 résume les moyennes des taux de chaque variable prosodique. Nous avons réalisé un test U de Mann-Whitney pour comparer les taux des pauses silencieuses, des pauses remplies et des allongements vocaliques des participantes BC à ceux des participantes HC. Les résultats ne montrent pas de différence significative entre les groupes pour ces variables prosodiques.

Contrôle des co-facteurs

Certains facteurs psychologiques ou états comme la fatigue peuvent influencer la production d'un discours. Nous avons réalisé une corrélation de Spearman pour examiner la relation entre les scores d'anxiété, de dépression et de fatigue et la durée des pauses silencieuses.

Les résultats ne montrent pas de corrélation significative entre les durées médiane et moyenne des pauses silencieuses et le score d'anxiété ($r(24) = .317$, $p = .114$, et $r(24) = .264$, $p = .192$), ni entre la durée médiane des pauses silencieuses et le score de dépression ($r(24) = .346$, $p = .084$). En revanche, les résultats suggèrent une corrélation significative entre les durées médiane et moyenne des pauses silencieuses et le score de fatigue ($r(24) = -.446$, $p = .022$ et $r(24) = - 473$, $p = .015$), ainsi que la durée moyenne des pauses silencieuses et le score de dépression ($r(24) = .458$, $p = .019$). Toutefois, aucun score n'ayant été considéré pathologique, ces résultats doivent être interprétés avec précaution.

DISCUSSION ET CONCLUSION

Ce travail décrit les différentes étapes nécessaires à la réalisation d'une analyse de la prosodie à travers une application concrète chez des participantes ayant eu un cancer du sein. La production de parole de treize participantes post-cancer du sein ayant un trouble cognitif induit par un cancer et ses traitements et celle de treize participantes contrôles saines appariées a été obtenue au moyen d'une tâche de narration élicitée par des images. Le corpus final, composé de vingt-six échantillons de parole pour une durée totale de trente-cinq minutes, a ensuite été analysé grâce à une méthode spécifiquement développée pour les besoins de l'étude.

La méthode employée utilise des outils informatiques automatisés, hors ligne et *open source*. Outre le gain de temps procuré par ces outils, leur utilisation protège le contrôle des données et de leur traitement. L'alternance entre SPPAS et Praat constitue toutefois une limite de cette méthode. En effet, elle nécessite la maîtrise de deux logiciels et peut allonger le temps d'analyse. Il est possible de réaliser l'ensemble de l'analyse prosodique sur SPPAS. Cependant, la visualisation du spectrogramme fourni par Praat rend l'ajustement des frontières plus aisé. De futures études veilleront à développer un outil combinant les fonctionnalités de SPPAS et la précision de Praat afin d'alléger davantage la procédure.

Cette méthode peut s'adapter à d'autres pathologies ou syndromes. Son application clinique peut être multiple. Elle permet par exemple d'établir une ligne de base et d'évaluer l'évolution des patients, de confirmer ou d'infirmer une intuition clinique en comparant la production d'un patient à des données normatives issues de la littérature, et/ou de proposer une analyse détaillée

TABLEAU 2 : Taux des pauses silencieuses, des pauses remplies et des allongements vocaliques.

Taux des variables prosodiques	BC *moyenne (ET)*	HC *moyenne (ET)*	*p*-value
Pauses silencieuses	0.270 (0.056)	0.224 (0.072)	0.101 ns
Pauses remplies	0.149 (0.062)	0.130 (0.072)	0.573 ns
Allongements vocaliques	0.172 (0.073)	0.161 (0.079)	0.880 ns

* = différence significative, ns = différence non-significative.

des difficultés linguistiques (recherche lexicale, programmation syntaxique, erreurs phonétiques, etc.).

Concernant l'application clinique que nous avons réalisée pour illustrer cette méthode, nos principaux résultats suggèrent que la durée des pauses d'hésitation silencieuses permet de différencier le groupe de participantes ayant eu un cancer du sein avec plainte cognitive des participantes contrôles, contrairement aux pauses remplies et aux allongements vocaliques. Ce résultat est similaire à ceux obtenus par Bóna (2024) qui a comparé les pauses dans le discours de personnes ayant des troubles cognitifs liés à une sclérose en plaques (SEP) à celles de participants sains lors de tâches de narration. Les pauses silencieuses des patients avec SEP sont plus fréquentes et plus longues que chez les contrôles. En revanche, les pauses remplies ne présentent pas de différence entre les deux groupes. Il est intéressant de noter que, bien que les pathologies sous-jacentes des deux populations (cancer et sclérose en plaques) diffèrent, les pauses silencieuses permettent d'identifier un TNCm chez des populations jeunes. Une explication possible serait que les pauses silencieuses, plus que les pauses remplies, soient le reflet d'une charge cognitive élevée.

Angelopoulou et al. (2024) soulignent la nécessité d'examiner séparément les types de pauses. En effet, chaque type de pauses répond à un besoin différent tel que des besoins linguistiques, interactionnels ou cognitifs. Par conséquent, elles peuvent refléter divers processus cognitifs. Il est donc nécessaire d'analyser isolément les pauses jouant un rôle de marqueur linguistique des TNCm.

Une des limites de notre étude est le petit effectif de participantes. Nos résultats sont donc à confirmer par une cohorte plus importante. Afin d'enrichir les connaissances sur les marqueurs de troubles cognitifs subtils en contexte de TNCm, de futures études pourront inclure d'autres paramètres prosodiques tels que la variabilité de F0 ou le débit de parole.

Pour conclure, notre étude montre l'intérêt des outils informatisés semi-automatiques pour l'évaluation de patients ayant un TNCm grâce au gain de temps et d'objectivité qu'ils procurent. Si l'utilisation de ces outils peut sembler à première vue fastidieuse, notre méthode détaillée assure la faisabilité technique d'une telle analyse prosodique en clinique.

Enfin, notre étude participe de l'évolution de la profession d'orthophoniste grâce à l'utilisation d'outils d'analyse automatisés ou d'intelligence artificielle qu'elle propose.

REMERCIEMENTS

Les auteurs remercient Farha Gheurbi et Marion Buchy d'avoir réalisé l'accord inter-annotateur de cette étude.

DÉCLARATION D'INTÉRETS

Les auteurs ont déclaré n'avoir aucun lien d'intérêt en relation avec cet article.

FINANCEMENTS

Amélie B. Richard est financée par une bourse doctorale attribuée par l'Université Montpellier Paul-Valéry et allouée par le ministère de l'Enseignement Supérieur, de la Recherche et de l'Innovation (MESRI).

RÉFÉRENCES

Amani, O., Mazaheri, M. A., Moghani, M. M., Zarani, F., & Choolabi, R. H. (2024). Chemotherapy-induced cognitive impairment in breast cancer survivors: A systematic review of studies from 2000 to 2021. *Cancer Reports, 7*(2), e1989. https://doi.org/10.1002/cnr2.1989

American Psychiatric Association. (2022). *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5-TR). https://doi.org/10.1176/appi.books.9780890425787

Angelopoulou, G., Kasselimis, D., Goutsos, D., & Potagas, C. (2024). A methodological approach to quantifying silent pauses, speech rate, and articulation rate across distinct narrative tasks: Introducing the Connected Speech Analysis Protocol (CSAP). *Brain Sciences, 14*(5), 466. https://doi.org/10.3390/brainsci14050466

Audacity Team (1999). *Audacity* (version 3.0) [logiciel informatique]. https://www.audacityteam.org/

Bézy, C., Renard, A., & Pariente, J. (dir.) (2016). *GRÉMOTS : batterie d'évaluation des troubles du langage dans les maladies neurodégénératives du GRECO*. De Boeck Supérieur.

Bigi, B. (2015). SPPAS - Multi-lingual approaches to the automatic annotation of speech. *The Phonetician. Journal of the International Society of Phonetic Sciences*, 111–112, 54–69. https://hal.science/hal-01417876

Boersma, P., & Weenink, D. (2024). *Praat: Doing phonetics by computer* (Version 6.1.55) [logiciel informatique]. https://www.fon.hum.uva.nl/praat/

Bóna, J. (2024). Pausing and fluency in speech of patients with relapsing-remitting multiple sclerosis. *Clinical Linguistics & Phonetics, 38*(4), 332–344. https://doi.org/10.1080/02699206.2023.2223347

Bonomi, A. E., Cella, D. F., Hahn, E. A., Bjordal, K., Sperner-Unterweger, B., Gangeri, L., Bergman, B., Willems-Groot, J., Hanquet, P., & Zittoun, R. (1996). Multilingual translation of the Functional Assessment of Cancer Therapy (FACT) quality of life measurement system. *Quality of Life Research, 5*(3), 309–320. https://doi.org/10.1007/BF00433915

Bryant, L., Spencer, E., & Ferguson, A. (2017). Clinical use of linguistic discourse analysis for the assessment of language in aphasia. *Aphasiology, 31*(10), 1105–1126. https://doi.org/10.1080/02687038.2016.1239013

Campione, E., & Véronis, J. (2002). *A large-scale multilingual study of silent pause duration.* Speech Prosody 2002, 11-13 avril. Aix-En-Provence. https://doi.org/10.21437/SpeechProsody.2002-35

Candea, M. (2000). *Contribution à l'étude des pauses silencieuses et des phénomènes dits "d'hésitation" en français oral spontané. Etude sur un corpus de récits en classe de français.* [thèse de doctorat, Université de la Sorbonne nouvelle - Paris III]. HAL. https://theses.hal.science/tel-00290143v1

Cella, D. F., Tulsky, D. S., Gray, G., Sarafian, B., Linn, E., Bonomi, A., Silberman, M., Yellen, S. B., Winicour, P., & Brannon, J. (1993). The Functional Assessment of Cancer Therapy scale: Development and validation of the general measure. *Journal of Clinical Oncology, 11*(3), 570–579. https://doi.org/10.1200/JCO.1993.11.3.570

Commission Nationale de l'Informatique et des Libertés - CNIL. (2016). *Le règlement général sur la protection des données—RGPD*. https://www.cnil.fr/fr/reglement-europeen-protection-donnees

Costa, D. S. J., Loh, V., Birney, D. P., Dhillon, H. M., Fardell, J. E., Gessler, D., & Vardy, J. L. (2018). The Structure of the FACT-Cog v3 in cancer patients, students, and older adults. *Journal of Pain and Symptom Management, 55*(4), 1173–1178. https://doi.org/10.1016/j.jpainsymman.2017.12.486

DiCanio, C. (2011). *Text replacement script for Praat* [logiciel informatique]. https://www.acsu.buffalo.edu/~cdicanio/scripts.html

Duez, D. (2001). Caractéristiques acoustiques et phonétiques des pauses remplies dans la conversation en français. *Travaux Interdisciplinaires du Laboratoire Parole et Langage (TIPA)*, 20, 31-48. https://hal.science/hal-00285404v1

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research, 12*(3), 189–198. https://doi.org/10.1016/0022-3956(75)90026-6

Goldman, J.-P., Thomas, F., Roekhaut, S., & Simon, A.-C. (2010). Étude statistique de la durée pausale dans différents styles de parole. *Actes des 28e Journées d'Étude sur la Parole (JEP)*, 161–164. http:// hdl.handle.net/2078.1/81909

Groupe de travail Traumatisme Crânio-Cérébral Léger (TCCL) en Île-de-France (2021). *Rapport du groupe de travail TCCL*. Agence Régionale de Santé Île-de-France. https://www.iledefrance.ars.sante.fr/traumatisme-cranio-cerebral-leger-tccl-en-ile-de-france-rapport-du-groupe-de-travail

Haute Autorité de Santé. (2018). *Patients présentant un trouble neurocognitif associé à la maladie d'Alzheimer ou à une maladie apparentée. Guide parcours de soins.* https://www.has-sante.fr/jcms/c_2906096/en/patients-presentant-un-trouble-neurocognitif-associe-a-la-maladie-d-alzheimer-ou-a-une-maladie-apparentee

He, R., Chapin, K., Al-Tamimi, J., Bel, N., Marquié, M., Rosende-Roca, M., Pytel, V., Tartari, J. P., Alegret, M., Sanabria, A., Ruiz, A., Boada, M., Valero, S., & Hinzen, W. (2023). Automated classification of cognitive decline and probable Alzheimer's dementia across multiple speech and language domains. *American Journal of Speech-Language Pathology, 32*(5), 2075–2086. https://doi.org/10.1044/2023_AJSLP-22-00403

Hommel, M., Miguel, S. T., Naegele, B., Gonnet, N., & Jaillard, A. (2009). Cognitive determinants of social functioning after a first ever mild to moderate stroke at vocational age. *Journal of Neurology, Neurosurgery & Psychiatry, 80*(8), 876–880. https://doi.org/10.1136/jnnp.2008.169672

Ivanova, O., Martínez-Nicolás, I., & Meilán, J. J. G. (2023). Speech changes in old age: Methodological considerations for speech-based discrimination of healthy ageing and Alzheimer's disease. *International Journal of Language & Communication Disorders, 59*(1), 13–37. https://doi.org/10.1111/1460-6984.12888

Janelsins, M. C., Kohli, S., Mohile, S. G., Usuki, K., Ahles, T. A., & Morrow, G. R. (2011). An update on cancer- and chemotherapy-related cognitive dysfunction: Current status. *Seminars in Oncology, 38*(3), 431–438. https://doi.org/doi:10.1053/j.seminoncol.2011.03.014

JASP Team. (2023). *JASP* (version 0.17.3) [logiciel informatique]. https://jasp-stats.org/team/

Jim, H. S. L., Phillips, K. M., Chait, S., Faul, L. A., Popa, M. A., Lee, Y.-H., Hussin, M. G., Jacobsen, P. B., & Small, B. J. (2012). Meta-analysis of cognitive functioning in breast cancer survivors previously treated with standard-dose chemotherapy. *Journal of Clinical Oncology, 30*(29), 3578–3587. https://doi.org/10.1200/JCO.2011.39.5640

Joly, F., Lange, M., Rigal, O., Correia, H., Giffard, B., Beaumont, J. L., Clisant, S., & Wagner, L. (2012). French version of the Functional Assessment of Cancer Therapy–Cognitive function (FACT-Cog) version 3. *Supportive Care in Cancer, 20*, 3297–3305. https://doi.org/10.1007/s00520-012-1439-2

Kassambara, A. (2016). *ggpubr: "ggplot2" Based publication ready plots* (version 0.6.0) [fichier de données]. https://doi.org/10.32614/CRAN.package.ggpubr

Lanzi, A. M., Saylor, A. K., Fromm, D., Liu, H., MacWhinney, B., & Cohen, M. L. (2023). DementiaBank: Theoretical rationale, protocol, and illustrative analyses. American *Journal of Speech-Language Pathology, 32*(2), 426–438. https://doi.org/10.1044/2022_AJSLP-22-00281

Mann, H. B., & Whitney, D. R. (1947). On a test of whether one of two random variables is stochastically larger than the other. *Annals of Mathematical Statistics, 18*(1), 50–60. https://doi.org/10.1214/aoms/1177730491

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

OSF (2011). Plateforme open source de soutien à la recherche. Center for Open Science. https://osf.io/sqwnz/?view_only=8947f20e707f476c9cf44a35ca9d7c32

Parjane, N., Cho, S., Ash, S., Cousins, K. A. Q., Shellikeri, S., Liberman, M., Shaw, L. M., Irwin, D. J., Grossman, M., & Nevler, N. (2021). Digital speech analysis in progressive supranuclear palsy and corticobasal syndromes. *Journal of Alzheimer's Disease, 82*(1), 33–45. Embase. https://doi.org/10.3233/jad-201132

Radford, A., Kim, J. W., Xu, T., Brockman, G., McLeavey, C., & Sutskever, I. (2022). Robust speech recognition via large-scale weak supervision. *arXiv: 2212.04356[*eess.AS*]*. https://doi.org/10.48550/ARXIV.2212.04356

Richard, A. B., Reilly, K. T., & Jacquin-Courtois, S. (2022, 17 juin). Que révèlent les disfluences sur le manque du mot rapporté par les patientes ayant un cancer du sein. *Actes des 34e Journées d'Etudes sur la Parole (JEP)*, 357-365. https://doi.org/10.21437/JEP.2022-38

Richard, A. B., Hirsch, F., Jacquin-Courtois, S., & Reilly, K. T. (2023). Self-interruptions in breast cancer patients who complain of anomia. Dans R. Skarnitzl & J. Volín (dir.), *Proceedings of the 20th International Congress of Phonetic Sciences* (p. 4012–4016). https://hal.science/hal-04195993v1

Richard, A. B., Foncelle, A., Hirsch, F., Jacquin-Courtois, S., Reilly, K., & Lelandais, M. (2024). *Speech markers of Cancer-Related Cognitive Impairment: A pilot study.* Speech Prosody 2024, 2-5 juillet. Leiden. https://doi.org/10.31219/osf.io/xz46h

RStudio, BPC. (2020). *RStudio: Integrated development for R* [logiciel informatique]. https://support.posit.co/hc/en-us/articles/206212048-Citing-RStudio

Spearman, C. (1904). The proof and measurement of association between two things. *The American Journal of Psychology, 15*(1), 72-101. https://doi.org/10.2307/1412159

Themistocleous, C., Eckerström, M., & Kokkinakis, D. (2020). Voice quality and speech fluency distinguish individuals with Mild Cognitive Impairment from healthy controls. *PLOS ONE, 15*(7), e0236009. https://doi.org/10.1371/journal.pone.0236009

Van Dyk, K., Crespi, C. M., Petersen, L., & Ganz, P. A. (2020). Identifying Cancer-Related Cognitive Impairment using the FACT-Cog perceived cognitive impairment. *JNCI Cancer Spectrum, 4*(1), pkz099. https://doi.org/10.1093/jncics/pkz099

Vincze, V., Szatlóczki, G., Tóth, L., Gosztolya, G., Pákáski, M., Hoffmann, I., & Kálmán, J. (2021). Telltale silence: Temporal speech parameters discriminate between prodromal dementia and mild Alzheimer's disease. *Clinical Linguistics & Phonetics, 35*(8), 727–742. https://doi.org/10.1080/02699206.2020.1827043

Wagner, L. I., Sweet, J., Butt, Z., Lai, J. S., & Cella, D. (2009). Measuring patient self-reported cognitive function: Development of the Functional Assessment of Cancer Therapy–cognitive function instrument. *The Journal of Supportive Oncology, 7*(6), W32–W39.

Wickham, H., Chang, W., Henry, L., Pedersen, T. L., Takahashi, K., Wilke, C., Woo, K., Yutani, H., Dunnington, D., & van den Brand, T. (2016). *ggplot2: Create elegant data visusalisations using the grammar of graphics*. https://ggplot2.tidyverse.org

Zigmond, A. S., & Snaith, R. P. (1983). The Hospital Anxiety and Depression scale. *Acta Psychiatrica Scandinavica, 67*(6), 361–370. https://doi.org/10.1111/j.1600-0447.1983.tb09716.x