Laboratoire : SCALab Sciences Cognitives et Sciences Affectives, UMR CNRS 9193

École doctorale : Sciences de l'Homme et de la Société

Université de Lille

Amorçage rythmique et production de la parole dans la maladie de Parkinson

Par **Jules Fumel**

Sous la direction du Pr. **Anahita Basirat**

Thèse de doctorat présentée en vue de l'obtention du grade de Docteur en Psychologie

Présentée et soutenue publiquement le 25 Juin 2026

Devant un jury composé de :

Delphine DELLACHERIE Professeure des Universités, Université de Lille	Présidente
Serge PINTO Directeur de recherche CNRS, Université Aix-Marseille	Rapporteur
Marie VILLAIN Maîtresse de conférences HDR, Sorbonne Université	Rapporteure
Maëva GARNIER Chargée de recherche CNRS HDR, Grenoble INP	Examinatrice
Cédric PATIN Maître de conférences, Université de Lille	Examinateur
Anahita BASIRAT Professeure des Universités, Université de Lille	Directrice de thèse

Remerciements

Après un peu plus de 3 ans et demi de thèse, voilà venu le moment de la conclure. Ces années, intenses, ont été d'une grande richesse, faites de nouvelles rencontres et du soutien précieux des personnes sur lesquelles je sais pouvoir compter. Avant d'entrer dans le vif du sujet, je vais profiter de ces quelques lignes pour remercier celles et ceux qui m'ont accompagné dans ce périple.

Je tiens tout d'abord à exprimer ma profonde gratitude à ma directrice de thèse, Anahita Basirat. Un grand merci, Anahita, de m'avoir fait confiance pour ce projet et de m'avoir guidé sur les chemins de la recherche. Tu as été d'un soutien constant, à la fois exigeant, bienveillant et profondément humain. Merci pour ta disponibilité, tes conseils experts et ta manière de guider sans jamais imposer. J'ai beaucoup appris à tes côtés, et cette thèse te doit évidemment énormément.

Je remercie également les membres de mon jury de thèse : Marie Villain, Serge Pinto, Delphine Dellacherie, Maëva Garnier et Cédric Patin, d'avoir accepté de prendre sur leur temps précieux pour lire ce travail et préparer nos échanges lors de la soutenance.

Mes remerciements vont également aux membres de mon CSI, Hélène Loevenbruck et Bruno Facon. Merci pour vos précieux conseils et votre bienveillance lors de nos entretiens.

Je pense aussi aux membres du labo SCAlab et en premier lieu ceux sans qui tout serait plus compliqué, pour ne pas dire impossible. Je vais bien sûr commencer par remercier Manu qui sait nous sortir des méandres administratifs comme personne, et toujours avec le sourire. Je pense aussi à Sabine qui a dû me rappeler à chaque déplacement comment fonctionnent les ordres de mission… Merci pour ta patience. Sur un registre un peu plus scientifique, un grand merci à Laurent Ott. Mes travaux te doivent beaucoup, ton efficacité est légendaire et tes explications mathématiques/statistiques sont limpides, même pour un orthophoniste.

Un grand merci également à l'ensemble de l'équipe Langage. Les temps d'échanges lors des séminaires ont toujours été stimulants et passionnants. L'ambiance qui règne dans l'équipe a rendu ces années d'autant plus agréables. Je remercie aussi l'ensemble des doctorants, toujours très accueillants. Un clin d'œil spécial à mes deux co-doctorants. Mélen, merci pour ta bonne humeur constante. Anaëlle, je suis très heureux d'avoir déjà pu participer à des conférences à tes côtés. En espérant que cela se reproduise à l'avenir.

Je souhaite exprimer ma reconnaissance à Antoine et Louison. Merci Antoine pour les figures, les outils et toutes ces magnifiques choses que tu sais faire en Python (et bientôt en Matlab ?). Merci Louison d'avoir été mon binôme sur Tourcoing. Ton expertise en EEG, autant qu'en blitz et en foodtrucks, ont égayé mon séjour à la Plaine Images.

Merci également à tous les étudiants rencontrés lors de mes enseignements. Un merci plus particulier à celles ayant participé plus directement aux études : Sarah, Clémence, Azénor, Lola, Lisa, Carla, Clara, Gwenaël, Maya, Anaïs, Apolline et Eloïse.

Je souhaite exprimer également ma gratitude à tous ceux qui m'ont accompagné dans la rédaction des articles. Je pense plus particulièrement à Paolo Mairano pour son aide sur les alignements phonétiques et à Leonardo pour les explications d'analyses sur Praat. Merci aux membres de la plateforme IrDive pour l'accueil et leur disponibilité. Ma reconnaissance va

également au personnel hospitalier du CHU de Lille qui m'a toujours facilité la tâche pour les passations dans leurs murs. Un grand merci au Pr Caroline Moreau et David Devos pour leur aide et leur accueil.

Je ne peux parler du CHU sans penser à mes patients et à mes formidables collègues de Swynghedauw. Merci à vous d'avoir compris quand je vous ai quittés pour 3 ans et des brouettes. Je suis heureux à l'idée de vous retrouver à la rentrée.

Avant de passer aux remerciements adressés à mon cercle proche, j'aimerais saluer et remercier infiniment Sylvie Donnat. Je ne vous remercierai jamais assez de la confiance que vous m'avez donnée lors de mes études égyptologiques. Je ne serais pas là où j'en suis aujourd'hui sans vous.

Hors des études, mes premières pensées vont à mes amis, qui ont su égayer ces années de travail acharné. Je ne vous ai pas vus autant que souhaité ces dernières années mais je compte bien me rattraper à l'avenir ! Je commence bien entendu par Rémi, Florence et leur petite Cléo. Merci d'être toujours à nos côtés. Hâte de reprendre nos longues sessions de jeu Rémi et merci Florence d'avoir mis un peu plus d'Alsace dans sa vie. J'ai une pensée particulière pour tes parents Rémi. Merci aussi à Isabelle, Baptiste et leur moyen Etienne. Merci Isabelle de toujours garder le sourire et de savoir nous faire relativiser. Merci à Capi, Vincent et à leurs enfants, Léo et Olivia. C'est toujours un bonheur de vous retrouver malgré les kilomètres qui nous séparent. Merci enfin à Hugo, Géraldine et leur petit Arthur. Nos rencontres rouennaises sont toujours géniales.

Je remercie également toute ma famille et belle-famille. Merci à ma sœur, Flore-Anne, et à sa famille Nicolas, Clémence et Maxence de toujours mettre les petits plats dans les grands. Merci à mon frère Félicien, à sa famille Marine, Lou et Sacha. Merci de me tenir informé des dernières nouveautés vidéoludiques, bien que cela puisse être parfois frustrant. Un merci infini à mes parents qui ont toujours tout fait pour que je puisse m'épanouir. Merci de m'avoir laissé vivre mes expériences et m'avoir appris à être curieux.

Je finis ces remerciements en ayant un mot pour les trois personnes les plus importantes de ma vie. A mes enfants, Soline et Baptiste tout d'abord. Vous ne pouvez pas savoir à quel point je me réjouis de pouvoir mettre derrière moi ce travail pour vous retrouver pleinement. Baptiste, je suis si fier de toi. C'est toujours surprenant comment, du haut de tes 5 ans, tu arrives à nous éclairer. J'ai hâte de pouvoir reprendre nos créations, nos sorties à vélo et qu'on se mette, bientôt, à notre orchestre. Soline, tu es un vrai rayon de lumière. Quoi qu'il arrive, tu as toujours ce grand sourire qui illumine tous ceux qui t'entourent. Ton arrivée parmi nous, en pleine thèse, a pimenté un peu plus notre quotidien mais les quelques heures de sommeil perdues compensent largement celles passées avec toi. J'ai hâte de pouvoir continuer à t'apprendre quelques pas de danse.

Enfin, je ne trouverai jamais assez de mots pour remercier celle qui m'accompagne. Merci pour tout, Alix. Ton soutien permanent, ta douceur, ton intelligence et tes rires sont des piliers essentiels à ma vie. Merci de m'avoir laissé m'engager dans cette aventure un peu folle du doctorat. Ces derniers mois ont parfois été difficiles, mais ta présence, ta patience et ta confiance m'ont aidé à garder le cap. Je mesure aussi tout ce que cette thèse t'a demandé, et je t'en remercie du fond du cœur. J'ai hâte que s'ouvre maintenant un temps plus doux, pour nous et pour notre famille.

Table des matières

TABLE DES MATIERES 4

RESUME 6

ABSTRACT 8

RESUME GRAND PUBLIC 10

AVANT-PROPOS 13

INTRODUCTION GENERALE 14

PREMIERE PARTIE CONTEXTE SCIENTIFIQUE ET PROBLEMATIQUE 17

CHAPITRE 1 MALADIE DE PARKINSON ET DYSARTHRIE HYPOKINETIQUE 18
1.1 DESCRIPTIF GENERAL 18
1.2 LA DYSARTHRIE HYPOKINETIQUE 21
1.3 PRISE EN CHARGE DES TROUBLES DE LA PAROLE DANS LA MALADIE DE PARKINSON 26
CHAPITRE 2 RYTHME, INTERVENTIONS ET AMORÇAGE RYTHMIQUE 31
2.1 INTERVENTIONS BASEES SUR LE RYTHME DANS LA MALADIE DE PARKINSON 31
2.2. L'AMORÇAGE RYTHMIQUE 38
CHAPITRE 3 MODELES ET CADRE INTERPRETATIF 46
3.1. UN CADRE INTEGRATIF MUSIQUE–PAROLE POUR PENSER L'AMORÇAGE RYTHMIQUE 47
3.2. MODELES DE PRODUCTION DE LA PAROLE 51
CHAPITRE 4 PROBLEMATIQUE ET OBJECTIFS 59

DEUXIEME PARTIE CONTRIBUTIONS ET ETUDES EXPERIMENTALES 64

CHAPITRE 1 MESURES ACOUSTIQUES DE LA PROSODIE DANS LA MALADIE DE PARKINSON : REVUE SYSTEMATIQUE ET META-ANALYSE BAYESIENNE 65
CHAPITRE 2 EFFETS DE L'AMORÇAGE RYTHMIQUE SUR LA PRODUCTION DE LA PAROLE DANS LA MALADIE DE PARKINSON ET L'ATAXIE SPINOCEREBELLEUSE 106
CHAPITRE 3 AU-DELA DU RYTHME 157
INTRODUCTION 157
METHODES 158
RESULTATS 160
DISCUSSION 166
CONCLUSION 173
CHAPITRE 4 DU RYTHME PERÇU A LA PAROLE PRODUITE : SYNCHRONISATION SENSORIMOTRICE ET AMORÇAGE RYTHMIQUE 175
INTRODUCTION 175
METHODES 179
RESULTATS 188
DISCUSSION 193
CONCLUSION 199
CHAPITRE 5 DE LA THEORIE NEUROCOGNITIVE A LA PRATIQUE ORTHOPHONIQUE 201

TROISIEME PARTIE DISCUSSION GENERALE 234

CHAPITRE 1 SYNTHESE TRANSVERSALE 235
CHAPITRE 2 LIMITES GENERALES ET REFLEXIONS METHODOLOGIQUES 248
CHAPITRE 3 PERSPECTIVES.....259

REFERENCES.....264

Résumé

Les troubles de la parole sont fréquents dans la maladie de Parkinson et correspondent le plus souvent à une dysarthrie hypokinétique. Parmi eux, les altérations prosodiques occupent une place importante, mais restent encore à préciser. Elles peuvent concerner la variabilité mélodique, le débit, l'organisation temporelle de la parole ou les pauses. Les traitements médicamenteux ne les résolvant pas, les approches fondées sur le rythme constituent une piste intéressante. Les stimulations auditives rythmiques sont déjà utilisées dans la maladie de Parkinson, mais leur intérêt pour la parole demeure moins établi. L'amorçage rythmique, qui consiste à exposer un individu à une structure rythmique avant une tâche, pourrait soutenir la perception temporelle, l'anticipation et la préparation motrice.

L'objectif général de cette thèse était d'examiner dans quelle mesure l'amorçage rythmique peut influencer la production de la parole, notamment dans la maladie de Parkinson, et d'interroger ces effets pour mieux comprendre les troubles prosodiques. Une première étude, sous la forme d'une revue systématique et d'une méta-analyse bayésienne, a permis de préciser les dimensions prosodiques les plus robustement altérées dans cette pathologie. Les résultats suggèrent une diminution de la variabilité de la fréquence fondamentale et une augmentation de la durée des pauses, tandis que d'autres mesures apparaissent plus hétérogènes.

Les études expérimentales ont ensuite évalué les effets d'amorces rythmiques auditives régulières et irrégulières sur plusieurs dimensions de la parole. Une première étude a été menée auprès de personnes atteintes de la maladie de Parkinson, de personnes atteintes d'ataxie spinocérébelleuse et de participants contrôles. Dans les trois groupes, l'effet le plus robuste concernait le temps d'initiation de la parole, avec une facilitation après amorce régulière. Les effets observés sur le débit, les métriques rythmiques, les pauses ou la variabilité mélodique étaient plus limités, et davantage dépendants de la population ou de la condition. Ces résultats suggèrent que l'amorçage rythmique agit surtout sur la préparation temporelle globale de la parole, plutôt que sur le contrôle fin de l'ensemble des paramètres prosodiques.

Une deuxième étude a été réalisée auprès de personnes sans pathologie neurologique. Elle visait à comparer différentes amorces régulières et irrégulières et à examiner si leurs effets sur la parole étaient associés aux compétences musicales et rythmiques des participants, évaluées par des tâches de synchronisation sensorimotrice. Cette étude a

confirmé l'effet bénéfique des amorces régulières sur le temps d'initiation de la parole, mais les compétences rythmiques mesurées ne semblent pas prédire la sensibilité à l'amorçage.

Enfin, une enquête menée auprès d'orthophonistes a interrogé l'usage clinique des interventions fondées sur la musique et le rythme. Elle met en évidence l'intérêt perçu de ces approches dans la pratique orthophonique, tout en soulignant la nécessité de mieux articuler preuves scientifiques, caractéristiques individuelles des patients, objectifs thérapeutiques et contraintes de terrain.

Dans leur ensemble, ces travaux suggèrent que l'amorçage rythmique auditif peut moduler certains aspects de la production de la parole, avec un effet particulièrement crédible sur le temps d'initiation. Ils invitent toutefois à considérer ces effets comme sélectifs, dépendants des tâches, des profils individuels et des caractéristiques des amorces. Cette thèse contribue ainsi à une meilleure compréhension des liens entre rythme, prosodie et contrôle moteur de la parole, tout en ouvrant des perspectives pour le développement d'approches orthophoniques fondées sur le rythme. Les travaux futurs devront préciser les paramètres optimaux des amorces, identifier les profils de patients les plus susceptibles d'en bénéficier et explorer les mécanismes sous-jacents à ces effets.

Abstract

Speech disorders are common in Parkinson's disease and most often take the form of hypokinetic dysarthria. Among these disorders, prosodic alterations play a significant role, but their precise nature remains to be clarified. They may involve melodic variability, speech rate, temporal organization of speech, or pauses. These dysarthric disorders are not resolved by standard drug treatments. In this context, rhythm-based approaches represent a particularly interesting avenue of investigation. Auditory rhythmic stimulation is already used in Parkinson's disease, particularly to support certain aspects of motor control such as walking. However, its benefits for speech remain less well established. Rhythmic priming, which involves exposing an individual to a rhythmic structure prior to performing a task, offers particularly interesting prospects in this regard. It could support temporal perception, anticipation and motor preparation, and thus modulate certain aspects of speech production.

The aim of this thesis was to examine the extent to which rhythmic priming can influence speech production, particularly in people with Parkinson's disease, and to investigate the effects observed in order to better understand and manage prosodic disorders. An initial study, in the form of a systematic review and a Bayesian meta-analysis, identified the prosodic dimensions most robustly altered in Parkinson's disease. The results suggest a reduction in fundamental frequency variability and an increase in pause duration. Other measures, however, appear more heterogeneous.

Experimental studies then assessed the effects of regular and irregular auditory rhythmic primes on various aspects of speech. An initial study was conducted with two groups of participants with neurodegenerative disorders —Parkinson's disease and spinocerebellar ataxia— and a control group. In these three groups, the most robust effect concerned speech initiation time, with facilitation following a regular prime. The effects observed on rate, rhythmic metrics, pauses or melodic variability were more limited, and more dependent on the population or condition. These results suggest that rhythmic priming acts primarily on the overall temporal preparation of speech, rather than on the fine and systematic control of all prosodic parameters.

A second study was conducted with individuals without neurological conditions. This study sought to compare different regular and irregular priming patterns and to link the observed effect of priming on speech with the participants' musical and rhythmic skills, in particular through assessment using sensorimotor synchronization tasks. This study confirmed

the beneficial effect of regular priming on speech initiation time. The rhythmic skills measured do not appear to predict the effects of rhythmic priming on speech.

Finally, a survey of speech and language therapists examined the clinical use of interventions based on music and rhythm. This study highlights the perceived value of these approaches in speech and language therapy practice, whilst emphasizing the need to better align scientific evidence, patients' individual characteristics, therapeutic objectives and practical constraints.

Taken together, these studies suggest that auditory rhythmic priming can modulate certain aspects of speech production, with a particularly credible effect on speech initiation time. However, they suggest that these effects should be regarded as selective, depending on the task, individual profiles and the characteristics of the primes. This thesis thus contributes to a better understanding of the links between rhythm, prosody and motor control of speech, whilst opening up prospects for the development of rhythm-based speech therapy approaches. Future research should, in particular, specify the optimal parameters of primes, identify the patient profiles most likely to benefit from them, and explore more directly the mechanisms underlying these effects.

Résumé Grand Public

Les troubles de la parole sont fréquents dans la maladie de Parkinson. Ils correspondent le plus souvent à ce que l'on appelle une dysarthrie, c'est-à-dire une difficulté à contrôler certains aspects de la parole. Ces troubles ne concernent pas seulement l'articulation ou la voix. Ils peuvent aussi toucher la prosodie, la « musique » de la parole : les variations de hauteur de la voix, le débit, l'organisation du rythme ou encore les pauses. Ces éléments jouent un rôle important dans l'expressivité de la parole et dans la fluidité des échanges.

Les traitements médicamenteux habituellement utilisés dans la maladie de Parkinson ne permettent pas toujours d'améliorer ces difficultés de parole. Dans ce contexte, le rythme constitue une piste intéressante. Les stimulations auditives rythmiques consistent à faire entendre un rythme régulier afin d'aider une personne à caler son mouvement sur ce rythme. Elles sont déjà utilisées dans la maladie de Parkinson, notamment pour soutenir efficacement la marche. Pour la parole, cependant, parler directement en synchronie avec un rythme pourrait modifier le caractère naturel de la production. Cette thèse s'intéresse donc à l'amorçage rythmique auditif : le rythme est entendu avant la production de parole, et non pendant celle-ci. Il pourrait ainsi aider à préparer l'action à venir, sans imposer de parler dans un rythme donné.

L'objectif général de cette thèse était donc d'examiner si l'amorçage rythmique auditif peut influencer la production de la parole, en particulier chez les personnes atteintes de la maladie de Parkinson. Il s'agissait aussi de mieux comprendre ce que ces effets peuvent apporter à l'étude des troubles touchant à la prosodie et, à plus long terme, à leur prise en charge.

Une première étude a consisté à faire une synthèse de la littérature scientifique sur la prosodie dans la maladie de Parkinson. Cette synthèse a permis d'identifier les aspects de la parole les plus clairement touchés. Les résultats montrent notamment que les personnes atteintes de la maladie de Parkinson présentent souvent une réduction des variations de hauteur de la voix, ainsi qu'un allongement des pauses. D'autres mesures de la prosodie apparaissent en revanche plus variables selon les études.

Les études expérimentales ont ensuite évalué les effets de différents types d'amorces rythmiques : des amorces régulières et des amorces irrégulières. Une première étude a été menée auprès des personnes atteintes de la maladie de Parkinson, des personnes atteintes

d’ataxie spinocérébelleuse (une pathologie neurodégénérative touchant le cervelet et impliquant également des troubles prosodiques) et des personnes sans pathologie neurologique. Le résultat le plus robuste concerne le temps nécessaire pour commencer à parler. Après l’écoute d’une amorce régulière, les participants avaient tendance à initier leur parole plus rapidement. En revanche, les effets sur d’autres aspects de la parole, comme le débit, des mesures rythmiques plus fines de la parole au niveau des voyelles ou des consonnes, les pauses ou les variations mélodiques de la voix, étaient plus limités et moins systématiques.

Une deuxième étude a été menée auprès de personnes sans pathologie neurologique. Elle visait à comparer différents rythmes réguliers et irréguliers, mais aussi à examiner si les compétences musicales ou rythmiques des participants pouvaient expliquer leur sensibilité à l’amorçage rythmique. Les résultats ont confirmé l’effet bénéfique du rythme régulier sur le temps d’initiation de la parole. En revanche, les compétences rythmiques mesurées, notamment à travers des tâches de synchronisation avec un rythme, ne semblaient pas prédire clairement l’effet de l’amorçage sur la parole.

Enfin, une enquête menée auprès d’orthophonistes a permis d’examiner la place de la musique et du rythme dans les pratiques cliniques. Cette enquête montre que ces approches suscitent un intérêt réel chez les professionnels, mais que leur utilisation reste variable. Elle souligne aussi la nécessité de mieux relier les connaissances scientifiques, les besoins individuels des patients, les objectifs thérapeutiques et les contraintes concrètes de la pratique.

Pris dans leur ensemble, ces travaux suggèrent que l’amorçage rythmique auditif peut influencer certaines dimensions de la parole, avec un effet particulièrement marqué sur le démarrage de la production orale. Ces effets doivent toutefois être considérés avec prudence. Ils ne concernent pas tous les paramètres de la parole de la même manière et semblent dépendre des tâches, des personnes et des caractéristiques des rythmes entendus. Cette thèse contribue ainsi à mieux comprendre les liens entre rythme, prosodie et contrôle moteur de la parole. Elle ouvre également des perspectives pour développer des approches orthophoniques fondées sur le rythme, à condition de mieux préciser quels rythmes utiliser, pour quels patients, et avec quels objectifs.

Avant-propos

Mes études d'orthophonie au sein du Centre de Formation Universitaire en Orthophonie de l'UFR3S ainsi que mon exercice clinique au sein du Centre de rééducation, de réadaptation et de Soins de suite Pierre Swynghedauw, m'ont sensibilisé aux troubles neurologiques, qu'ils soient acquis ou neurodégénératifs. J'ai notamment eu l'occasion de constater combien ces deux situations critiques diffèrent pour les patients, et leurs proches, entre l'éclair dans un ciel serein et un diagnostic dont ils connaissent la finalité, tout en ignorant le chemin qui y mènera. Parmi la constellation de troubles qu'impliquent ces pathologies, les aspects langagiers et de parole m'ont toujours particulièrement intéressé, et plus encore la prosodie, pour ce qu'elle partage entre musique et langage. Lors de mon exercice professionnel, j'ai pu constater que les patients étaient particulièrement friands des méthodes de rééducation basés sur le rythme et la musique, en solo comme en groupe. Pourtant, si cette utilisation est clairement établie dans la pratique clinique, le manque de données probantes et scientifiquement fondées m'interrogeait. Cela sonnait donc comme une évidence de me lancer dans ce projet de thèse, que je vois comme la possibilité de contribuer, modestement, à mieux comprendre le rôle que peuvent jouer le rythme et la musique d'un point de vue rééducatif sur la production de la parole, à en préciser les fondements neurocognitifs et à renforcer le lien entre recherche et pratique clinique.

Cette thèse a été réalisée au sein du laboratoire SCALab, UMR 9193 à l'Université de Lille. Elle a été rendue possible grâce aux financements issus du projet ANR EntrainPark ainsi que par l'obtention d'une bourse France Parkinson. Une partie du recueil des données a été réalisée à la Fédération de Recherche Sciences et Cultures du Visuel (FR CNRS 2052 SCV).

Introduction générale

Les troubles de la parole font partie des manifestations fréquentes de la maladie de Parkinson (MP) et compliquent la communication au quotidien, modifiant le caractère naturel de la parole et contribuant de manière importante à la réduction des interactions sociales (Atkinson-Clement et al., 2019). Définis comme dysarthrie hypokinétique, ces troubles ne se limitent pas à une baisse de l'intensité vocale ou à une moindre précision articulatoire. Ils concernent plus largement plusieurs dimensions de la production verbale, parmi lesquelles la prosodie et l'organisation temporelle occupent une place centrale (Duffy, 2020).

Ces dimensions présentent un intérêt particulier, à la fois sur le plan clinique et sur le plan scientifique. Sur le plan clinique, elles participent directement à l'efficacité de la communication : une parole monotone, sans relief ou mal segmentée peut rapidement devenir coûteuse pour le locuteur comme pour l'interlocuteur. Sur le plan scientifique, elles renvoient à des mécanismes complexes, qui engagent conjointement l'exécution motrice, la planification, le séquençage et l'organisation temporelle de la parole. Dans la MP, toutes les composantes ne semblent pas atteintes avec la même constance. Certaines altérations des paramètres acoustiques de la prosodie sont rapportées de façon relativement robuste dans la littérature, tandis que d'autres résultats apparaissent plus variables selon les études, les tâches ou les mesures retenues (p. ex. Frota et al., 2021). Cette hétérogénéité invite à préciser plus finement quelles dimensions de la parole sont les plus affectées, et à quel niveau elles doivent être étudiées.

Dans ce contexte, la question du rythme apparaît particulièrement intéressante. La parole est en effet une activité structurée dans le temps (Turk & Shattuck-Hufnagel, 2020). Produire un énoncé ne consiste pas seulement à enchaîner des segments linguistiques ; cela suppose aussi de les organiser selon une dynamique temporelle précise, faite d'alternances, de regroupements, d'accents et de pauses qui participent à la forme même du message. En parallèle, de nombreux travaux ont montré qu'une structure temporelle régulière peut orienter l'attention dans le temps, favoriser l'anticipation et préparer l'action (Large & Jones, 1999 ; Jones, 2019). Les liens entre musique, rythme et langage ont ainsi suscité un intérêt croissant, en particulier parce qu'ils semblent engager, au moins en partie, des mécanismes communs de traitement auditif, de synchronisation/entrainement neuronal et de couplage sensorimoteur (Fiveash et al., 2021). Appliquée à la production de la parole, cette perspective

conduit à s'interroger sur les effets qu'une structure rythmique entendue avant la production pourrait avoir sur la parole qui suit.

Une telle question paraît particulièrement pertinente dans la MP. En effet, les troubles de la parole observés dans cette pathologie conduisent à interroger les mécanismes de timing, de séquençage et d'initiation impliqués dans la parole. Ces mécanismes peuvent être mis en relation, d'une part, avec le rôle des ganglions de la base dans le traitement temporel et la sélection des programmes moteurs, et d'autre part, avec des cadres neurocomputationnels de la production de la parole, tels que DIVA/GODIVA, qui articulent planification, initiation, séquençage et contrôle sensorimoteur (Bohland et al., 2010 ; Manes et al., 2024 ; Tourville & Guenther, 2011). Dans cette perspective, l'amorçage rythmique pourrait constituer un outil d'étude particulièrement intéressant : il permet d'examiner dans quelle mesure un indice temporel externe peut influencer certains paramètres de la production orale. Il offre également un moyen d'étudier les relations entre perception du rythme, préparation motrice et production de la parole.

La présente thèse s'inscrit dans cette perspective générale. Elle porte sur l'effet de l'amorçage rythmique sur la production de la parole, principalement dans la MP. Son objectif n'est pas seulement de déterminer si un rythme entendu avant la parole influence la production ; il est aussi de mieux comprendre quelles dimensions de la parole sont concernées, quelles conditions produisent ces effets, et quelle peut être leur portée pour la compréhension des troubles de la parole et, plus largement, pour la réflexion clinique en orthophonie. Ce travail se situe ainsi au croisement de plusieurs champs d'étude : la phonétique clinique, la psychologie et les neurosciences cognitives, dont la recherche sur les relations entre musique et langage, et l'orthophonie.

Cet objectif général a conduit à articuler plusieurs niveaux d'analyse. Un premier axe a été de mieux cerner, à partir de la littérature et au travers d'une méta-analyse, les dimensions de la parole liées à l'expression de la prosodie les plus altérées dans la MP. Un deuxième axe a été d'examiner expérimentalement les effets de l'amorçage rythmique sur la parole, aussi bien en population générale que dans des populations cliniques. Cet axe s'est appuyé sur l'analyse acoustique de la parole produite après l'écoute d'amorces rythmiques, afin d'évaluer leurs effets sur différentes dimensions temporelles et prosodiques de la parole. Enfin, un dernier axe a été d'analyser la portée clinique de ces travaux. Cette réflexion s'appuie d'une part sur une enquête par questionnaire menée auprès d'orthophonistes concernant leurs

usages de la musique et du rythme, et d'autre part sur les perspectives ouvertes par ce travail, dans une logique de pratique fondée sur les preuves (Greenwell & Walsh, 2021).

Ce manuscrit est ainsi organisé en trois parties. La première présente le contexte théorique et la problématique générale de la thèse. Elle débute par la présentation de la MP et des troubles de la parole associés. Sont ensuite développées les interventions fondées sur le rythme dans cette pathologie, et au-delà, avec un focus sur l'amorçage rythmique. Enfin, les principaux cadres théoriques de ce travail, en lien avec le rythme et la production de la parole, sont exposés. La deuxième partie rassemble trois articles, publiés ou en révision, constituant le cœur du manuscrit : une méta-analyse des altérations prosodiques dans la MP, une étude sur les effets de l'amorçage rythmique sur l'organisation temporelle de la parole dans la MP et dans les ataxies spinocérébelleuses (ASC) et un article portant sur les pratiques des orthophonistes concernant l'usage de la musique et du rythme. L'étude sur l'amorçage rythmique est enrichie d'une partie complétant l'étude de l'organisation temporelle en s'intéressant à des éléments en lien avec la variabilité de la fréquence fondamentale (F0) et les pauses. Enfin, un chapitre est dédié à l'étude réalisée auprès de participants, sans troubles neurologiques, étude qui se propose de lier mesures acoustiques et synchronisation sensorimotrice. La troisième partie est consacrée à la discussion générale et aux perspectives, avec pour objectif de mettre en relation les résultats des différents travaux, d'en discuter les implications théoriques et cliniques et de se projeter dans une étude future.

Première partie

Contexte scientifique et problématique

Chapitre 1
Maladie de Parkinson et dysarthrie hypokinétique

1.1 Descriptif général

Longtemps décrite avant tout à partir de ses manifestations motrices cardinales, la MP est aujourd'hui reconnue comme une affection neurodégénérative aux retentissements bien plus larges. Si la bradykinésie, la rigidité et le tremblement de repos en constituent les signes les plus caractéristiques, la maladie affecte en réalité un ensemble de fonctions complexes qui ne se laissent pas réduire au seul champ du mouvement (Defebvre & Vérin, 2025). Parmi ces fonctions, la communication occupe une place particulière, car elle mobilise à la fois la parole, le langage, l'expressivité, l'ajustement à l'interlocuteur et, plus largement, des mécanismes d'initiation, de coordination sensorimotrice et de contrôle temporel. Cette perspective justifie que la MP soit abordée ici non seulement comme une pathologie du mouvement, mais aussi comme un cadre clinique dans lequel la parole et la communication représentent des enjeux majeurs. Des synthèses récentes soulignent d'ailleurs que les changements communicationnels dans la MP sont plus larges et bien plus variés que la voix hypophonique caractéristique, et qu'ils impliquent différents aspects de la production de la parole et des processus langagiers (Barnish et al., 2017 ; Basirat, 2025 ; Miller, 2017 ; Moreau & Pinto, 2019).

Sur le plan épidémiologique, la MP représente un enjeu majeur de santé publique. À l'échelle mondiale, la prévalence de la MP augmente nettement avec l'âge et semble avoir progressé au cours des dernières décennies. Une méta-analyse récente, portant sur les études publiées entre 1980 et 2023, estime la prévalence mondiale de la MP à 1,51 cas pour 1 000 personnes, avec des valeurs plus élevées dans les études les plus récentes, notamment 3,81 pour 1 000 pour la période 2010–2023. Cette prévalence apparaît également légèrement plus élevée chez les hommes et varie selon les régions et les indicateurs sociodémographiques des pays (Zhu et al., 2024). En France, la MP constitue la deuxième maladie neurodégénérative après la maladie d'Alzheimer et les maladies apparentées. Au 31 décembre 2020, environ

175 000 personnes étaient traitées pour une MP, soit une prévalence de 2,63 pour 1 000 habitants, et près de 26 000 personnes étaient nouvellement traitées au cours de l'année 2020 (Santé publique France, 2023). Sur le plan neuropathologique, la MP est principalement associée à une dégénérescence des neurones dopaminergiques de la substance noire pars compacta, structure intégrée aux circuits des ganglions de la base (Hornykiewicz, 2006). Ces circuits jouent un rôle central dans la sélection, l'initiation et l'ajustement de l'action. Leur implication ne se limite toutefois pas au seul contrôle moteur : elle concerne plus largement l'organisation temporelle du comportement, l'apprentissage procédural, certaines dimensions de la prise de décision et la planification de séquences complexes (Bloem et al., 2021 ; Defebvre & Vérin, 2025). Cette conception plus large est particulièrement pertinente lorsqu'il s'agit d'aborder la parole et la communication, qui reposent précisément sur l'articulation entre mouvement, temporalité, contrôle sensorimoteur et adaptation interactionnelle.

Les premiers signes cliniques apparaissent généralement lorsque la dégénérescence dopaminergique est déjà avancée et que les mécanismes compensatoires ne suffisent plus à maintenir un fonctionnement satisfaisant (Armstrong & Okun, 2020 ; Fearnley & Lees, 1991). La symptomatologie associe des manifestations motrices et non motrices, dont l'intrication contribue à l'hétérogénéité clinique de la maladie. Les signes non moteurs comprennent notamment des troubles du sommeil, des manifestations dysautonomiques, des symptômes sensitivo-douloureux, des troubles cognitifs et des troubles psycho-affectifs ; certains peuvent précéder l'apparition des symptômes moteurs (Bloem et al., 2021). La présentation clinique classique demeure néanmoins structurée autour du syndrome parkinsonien, associant la triade associant bradykinésie, rigidité et/ou tremblement de repos (Defebvre & Vérin, 2025). Cette triade reste indispensable à la description médicale de la MP, mais elle ne suffit pas à rendre compte de l'ensemble de ses conséquences fonctionnelles. Chez les personnes atteintes de la MP, les difficultés observées ne relèvent pas seulement d'une atteinte motrice générale. Elles concernent aussi des activités complexes, qui exigent une coordination fine, une initiation efficace, une temporalité d'exécution adaptée et un ajustement constant à l'environnement et au contexte.

La communication fait pleinement partie de ces activités. Elle mobilise la production motrice de la parole, bien sûr, mais aussi l'organisation du message, l'adaptation à l'interlocuteur, l'expressivité et la gestion dynamique de l'échange. Les troubles de la communication observés dans la MP ne peuvent donc pas être considérés comme un simple retentissement secondaire des symptômes moteurs ; ils constituent une dimension clinique à part entière, au croisement entre atteintes motrices, cognitives, affectives et interactionnelles (Miller, 2017 ; Prenger et al., 2020). Les conséquences de ces troubles sont importantes. Elles affectent la participation sociale, les relations interpersonnelles et, plus largement, le vécu quotidien des personnes concernées (Miller et al., 2006 ; Schalling et al. 2017, Yorkston et al. 2017) et de leurs proches (Baylor et al., 2024). Dans la vie quotidienne, elles peuvent se traduire par une difficulté accrue à se faire comprendre, à soutenir des échanges spontanés, à ajuster son discours à l'interlocuteur et à la situation ou à transmettre certaines nuances pragmatiques et émotionnelles. Il ne s'agit pas d'un trouble unique, mais d'un ensemble de limitations susceptibles de s'additionner : trouble moteur de la parole, troubles langagiers de production comme de compréhension, altérations de l'expressivité vocale et faciale, fluctuations cognitives et attentionnelles, ralentissement du traitement de l'information, voire diminution de l'initiative communicationnelle pragmatique (Altmann & Troche, 2011 ; Miller, 2017 ; de Souza & Beber, 2026). Les travaux de Barnish et al. (2016) illustrent cela, indiquant que, si l'intelligibilité contribue aux difficultés de communication quotidienne dans la MP, l'intervention d'autres dimensions, notamment cognitives et interactionnelles, est également importante.

Parmi toutes les composantes de ces difficultés communicationnelles, les troubles de la parole occupent une place centrale. Chez les personnes vivant avec la MP, la dysarthrie hypokinétique constitue le trouble moteur de la parole le plus classiquement décrit. Les changements de parole, de voix et de communication sont fréquemment rapportés par les personnes concernées, avec des estimations allant d'environ 70 % pour les altérations de parole et de voix à plus de 90 % lorsqu'un ensemble plus large de symptômes communicationnels est pris en compte (Hartelius & Svensson, 1994 ; Schalling et al., 2017). Ces troubles peuvent apparaître précocement au cours de l'évolution de la maladie (Rusz et al., 2011), au point d'être envisagés comme des marqueurs potentiels de diagnostic et de suivi

(Rusz et al., 2021a ; Rusz et al., 2024). Leur importance clinique tient non seulement à leur fréquence, mais aussi au fait qu'ils touchent une fonction omniprésente dans la vie quotidienne, mobilisée dans la quasi-totalité des interactions sociales, familiales et professionnelles. Dans la suite de ce chapitre, l'analyse sera centrée sur les troubles de la parole dans la MP, et plus particulièrement sur la dysarthrie hypokinétique et ses dimensions prosodiques.

1.2 La dysarthrie hypokinétique

1.2.1. Une description perceptive

La dysarthrie correspond à un trouble de l'exécution motrice de la parole, consécutif à une atteinte du système nerveux central ou périphérique, affectant les différents sous-systèmes impliqués dans la production de la parole, tant au niveau segmental que suprasegmental (Duffy, 2020). L'étymologie du terme peut toutefois induire en erreur si elle laisse entendre un trouble essentiellement articulatoire. En pratique, la dysarthrie concerne l'ensemble des sous-systèmes de la parole : respiration, phonation, résonance et articulation. À ce titre, le terme de « dysarthropneumophonie » a parfois été utilisé afin de souligner l'imbrication fréquente des atteintes articulatoires, respiratoires et phonatoires (Moreau et al., 2011). La classification clinique actuelle des dysarthries découle largement des travaux fondateurs de Darley et de l'équipe de la Mayo Clinic à la fin des années 1960 (Darley et al., 1969a, 1969b, 1975). À partir d'une analyse perceptive réalisée sur une tâche de lecture, ces auteurs ont caractérisé la parole dysarthrique selon 38 paramètres couvrant la respiration, la phonation, la résonance, l'articulation et les dimensions prosodiques. Cette démarche présente des limites connues, notamment l'évaluation sur une tâche unique, sans parole spontanée, ainsi que la description parfois incomplète de certains paramètres cliniques. Elle reste néanmoins structurante dans la manière de décrire et de penser les troubles moteurs de la parole.

Dans la MP, le profil classiquement décrit correspond à la dysarthrie hypokinétique. Sur le plan perceptif, celle-ci se caractérise par une réduction de la variation de hauteur, une réduction de l'accentuation, une monotonie d'intensité, une imprécision consonantique, des

pauses inappropriées, des accélérations brèves, une voix rauque ou soufflée, une baisse de la hauteur moyenne, ainsi que des modifications du débit de parole (Atalar et al., 2023 ; Brabenec et al., 2017 ; Darley et al., 1969a, 1969b, 1975 ; Pinto et al., 2010). Il se dégage ainsi un profil marqué par une atténuation des contrastes acoustiques et prosodiques. Cette dysarthrie ne peut donc être réduite à une simple addition de symptômes isolés, ni à la seule hypophonie ou à une imprécision segmentale. Elle affecte plus largement l'organisation de la parole comme activité motrice coordonnée, dans laquelle les variations de hauteur, d'intensité, de durée, de débit et de précision articulatoire contribuent à la structuration du message. Ses conséquences fonctionnelles ne concernent donc pas uniquement la qualité vocale, mais aussi l'intelligibilité, la transmission des indices prosodiques, l'expressivité et, plus largement, l'efficacité de la communication (Miller, 2017 ; Pinto et al., 2010).

À partir de ces éléments, plusieurs auteurs soulignent l'intérêt de distinguer, au sein de la dysarthrie hypokinétique, les dimensions principalement vocales, telles que l'intensité ou la qualité vocale, et les dimensions plus spécifiquement prosodiques, telles que l'intonation, l'accentuation, le rythme ou les variations temporelles de la parole (Pinto et al., 2010 ; Rusz et al., 2021b). Ces dimensions sont souvent associées, mais elles ne sont pas strictement superposables. Une voix faible, rauque ou soufflée ne suffit pas à résumer l'ensemble des altérations du rythme, de l'accentuation ou de l'intonation ; inversement, une parole monotone ne se réduit pas nécessairement à une baisse d'intensité. Cette distinction est cliniquement importante, car elle suggère que l'évaluation et la prise en soin ne peuvent pas se limiter à un travail vocal général : la prosodie peut constituer une cible spécifique à explorer et, le cas échéant, à cibler dans l'intervention (Atalar et al., 2023 ; Miller, 2017 ; Pinto et al., 2016).

1.2.2. Prosodie et dysprosodie

La prosodie désigne l'ensemble des phénomènes suprasegmentaux qui structurent la parole, notamment l'intonation, l'accentuation et le rythme (Brin-Henry et al., 2011). Si elle peut être appréhendée à partir de marqueurs acoustiques objectivables, elle renvoie également à des dimensions phonologiques et fonctionnelles plus abstraites. Elle contribue à l'organisation et à la hiérarchisation de l'information, au marquage des frontières prosodiques

et de la modalité de phrase, tout en véhiculant des indices paralinguistiques relatifs à l'attitude ou à l'état émotionnel du locuteur (Di Cristo, 2004, 2013 ; Lacheret, 2011). Peuvent donc être distinguées, une prosodie à dominante linguistique, impliquée dans l'organisation du message, et une prosodie à dominante émotionnelle, davantage liée à l'expression affective (Coulombe et al., 2023 ; Schwartz & Pell, 2012). Ces deux dimensions peuvent être altérées dans certaines pathologies d'origine neurologique, notamment après lésion de l'hémisphère droit ou plus largement après lésion cérébrale acquise (Benedetti et al., 2022 ; Stockbridge et al., 2022 ; Weed & Fusaroli, 2020).

La prosodie intervient également dans la gestion de l'interaction, en participant à l'orientation de l'attention, à l'organisation des tours de parole et, plus largement, à l'ajustement à l'interlocuteur (Bögel & Torreira, 2021). D'un point de vue descriptif, on distingue classiquement le rythme, l'intonation et l'accentuation, qui renvoient respectivement à l'organisation temporelle du signal, aux variations de hauteur et à la mise en relief de certains éléments du message (Di Cristo, 2013). Ces dimensions peuvent être étudiées à partir de paramètres acoustiques tels que la durée, la F0 et l'intensité (Di Cristo, 2013, Weed & Fusaroli, 2020). La durée renseigne sur le tempo, le débit et les pauses ; la F0, associée à la hauteur perçue, dépend de la fréquence de vibration des plis vocaux ; l'intensité reflète, elle, l'énergie du signal.

Parmi ces composantes, le rythme occupe une place centrale dans la présente thèse. Dans le domaine de la parole, il renvoie à l'organisation temporelle des unités prosodiques et à la distribution des proéminences au sein de l'énoncé, distribution qui contribue à structurer la parole en groupes perceptivement saillants (Di Cristo, 2016). Les syllabes proéminentes peuvent ainsi se réaliser comme des proéminences prosodiques, c'est-à-dire comme des unités qui se détachent de leur environnement sonore et qui peuvent être appréhendées à partir de paramètres tels que la durée syllabique, l'intensité ou la F0 (Di Cristo, 2016 ; Simon et al., 2008). En français, l'organisation prosodique repose sur plusieurs dimensions étroitement articulées, notamment l'accentuation, le phrasé et l'intonation. L'accentuation est principalement post-lexicale, avec un accent primaire final à valeur démarcative et un accent initial plus optionnel, susceptible, dans certains contextes, de contribuer à l'équilibre rythmique des groupes prosodiques (Di Cristo, 2013). Ces groupes s'inscrivent dans une organisation hiérarchique comprenant notamment le groupe accentuel, la phrase

intermédiaire et la phrase intonative (Astésano & Bertrand, 2016 ; Delais-Roussarie et al., 2015 ; Jun & Fougeron, 2000, 2002).

Dans ce cadre, la dysprosodie peut être comprise comme une altération des propriétés suprasegmentales de la parole, c'est-à-dire de l'organisation prosodique de l'énoncé. Dans la MP, elle renvoie à des altérations de plusieurs dimensions prosodiques, notamment de la variation intonative, du phrasé prosodique et, plus largement, de certaines dimensions temporelles et rythmiques, avec des profils qui varient selon les langues, les tâches et les niveaux d'analyse (Brabenec et al., 2017 ; Frota et al., 2021). Elle ne se réduit donc pas à une simple impression de parole monotone au sens courant du terme, mais peut également affecter l'organisation prosodique de l'énoncé, notamment le marquage des frontières, la structuration en groupes prosodiques et, selon les cas, la gestion des pauses. Les conséquences de ces atteintes ont été mises en évidence à la fois en production et en perception. Du côté de la production, la prosodie des personnes atteintes de la MP est souvent jugée moins expressive, moins informative et moins forte sur le plan émotionnel, y compris à des stades relativement précoces de la maladie (Pell et al., 2006). Du côté de la perception des indices prosodiques chez les personnes atteintes de la MP, les résultats apparaissent plus hétérogènes : plusieurs travaux ont rapporté des difficultés dans l'identification de significations linguistiques ou affectives (ex. joie) portées par la prosodie, tandis que d'autres suggèrent une relative préservation de certains traitements plus élémentaires selon les tâches considérées (Breitenstein et al., 2001 ; Lloyd, 1999 ; Pell, 1996). Des travaux plus récents suggèrent en outre que certaines dimensions prosodiques peuvent être davantage affectées en production qu'en perception, comme cela a été observé pour le marquage du contraste focal en mandarin : les personnes atteintes de la MP produisaient moins efficacement l'accent contrastif, tandis que leur perception de ces contrastes apparaissait relativement préservée (Chen & Sidtis, 2024). Ces altérations n'ont pas seulement une portée descriptive : elles peuvent avoir des conséquences concrètes sur l'efficacité de la communication, en perturbant la segmentation en unités de sens, en réduisant la clarté du marquage prosodique et en atténuant l'expression de l'attitude et de l'émotion. Cela est d'autant plus important que des données récentes suggèrent qu'un aplatissement de la F0 peut augmenter l'effort d'écoute et dégrader l'intelligibilité évaluée subjectivement, même en l'absence d'effet net sur les performances de transcription (Van Der Bruggen et al., 2024).

Dans une pathologie où les difficultés de communication reposent déjà sur une vulnérabilité multifactorielle, la dysprosodie apparaît ainsi comme une composante majeure. Cette importance clinique ne signifie toutefois pas que sa caractérisation soit aisée. La prosodie renvoie en effet à un ensemble de dimensions qui peuvent être étudiées à des niveaux distincts (perceptif, acoustique, phonologique ou fonctionnel), et dont la description dépend étroitement des tâches mobilisées, des paramètres retenus et de la langue étudiée. Cette complexité contribue probablement au fait que la dysprosodie, bien que fréquemment décrite cliniquement dans la MP, demeure parfois plus difficile à caractériser avec précision que d'autres aspects de la dysarthrie.

La dysprosodie décrite perceptivement dans la dysarthrie hypokinétique peut également être explorée de manière quantitative par l'analyse acoustique. Les outils disponibles facilitent l'extraction de mesures liées à la F0, à l'intensité et à la durée, permettant de préciser plus finement le profil des troubles et d'en suivre l'évolution (Rusz et al., 2024 ; Trouville et al., 2021). Au-delà de cette facilitation, elle permet également de rendre compte de réalité physique échappant à l'oreille humaine. Cette approche présente l'avantage de fournir des indices quantifiables, utilisables en clinique, dans des outils d'évaluation et de suivi, et plus largement dans la comparaison entre études.

Cette objectivation n'est toutefois pas sans difficulté. Les résultats de la littérature demeurent parfois hétérogènes selon les tâches, les méthodes d'extraction, les variables retenues et les caractéristiques des populations étudiées. Même pour des mesures centrales comme la variabilité de la F0, les conclusions ne sont pas toujours convergentes : certaines études rapportent une diminution de cette variabilité (Galaz et al., 2016), tandis que d'autres soulignent que les effets acoustiques liés à la MP peuvent dépendre fortement des tâches, des contextes prosodiques et des paramètres considérés (Cheang & Pell, 2007). Cette hétérogénéité ne remet pas en cause l'intérêt des mesures acoustiques ; elle souligne plutôt la nécessité d'identifier les marqueurs les plus robustes et les plus informatifs (Brabenec et al., 2017 ; Atalar et al., 2023). Au-delà de leur intérêt descriptif, ces mesures peuvent également contribuer à éclairer les mécanismes impliqués dans les altérations prosodiques observées dans la MP, notamment en lien avec le contrôle sensorimoteur et les circuits cortico-sous-corticaux (Brabenec et al., 2017 ; Manes et al., 2024). Leur interprétation

mécanistique reste toutefois encore partielle, ce qui renforce la nécessité de disposer d'indicateurs acoustiques fiables.

Avant de pouvoir discuter les mécanismes impliqués dans la dysprosodie parkinsonienne, il est nécessaire d'identifier les dimensions prosodiques les plus systématiquement altérées chez les personnes atteintes de la MP. Cette clarification constitue également un enjeu clinique, dans la mesure où elle permet de dégager des cibles prioritaires pour l'évaluation, le suivi et l'étude d'éventuelles interventions sur la parole. C'est dans cette optique qu'une revue systématique et une méta-analyse ont été menées dans cette thèse afin de préciser la fiabilité des altérations prosodiques rapportées dans la littérature (voir chapitre 1, deuxième partie).

1.3 Prise en charge des troubles de la parole dans la maladie de Parkinson

Lorsqu'on s'intéresse spécifiquement à la parole dans la MP, un constat s'impose d'emblée : l'amélioration des symptômes moteurs généraux ne s'accompagne pas nécessairement d'une amélioration parallèle des troubles de la parole (Fabbri et al., 2017 ; Tykalova et al., 2022). Cette dissociation est importante, car elle rappelle que la prise en soin de la communication ne peut pas être pensée comme un simple prolongement des traitements moteurs usuels. Sur le plan médical, la prise en charge de la MP repose encore principalement sur des traitements symptomatiques. Les approches dopaminergiques, en particulier la lévodopa, permettent souvent d'améliorer les manifestations motrices, surtout aux premiers stades de la maladie, mais elles ne modifient pas directement l'évolution neurodégénérative (Armstrong & Okun, 2020). A mesure que la pathologie progresse, la réponse à la médication devient toutefois moins stable, notamment en lien avec l'apparition de fluctuations motrices et de complications du traitement (Bloem et al., 2021).

La situation est différente lorsqu'on considère la parole et la voix. Les effets des traitements dopaminergiques sur ces dimensions apparaissent globalement modestes, irréguliers, et souvent insuffisants pour produire un bénéfice fonctionnel clair sur la

communication quotidienne. Plusieurs travaux anciens et récents convergent sur ce point : certains patients montrent des modifications mesurables sous traitement, mais celles-ci restent variables selon les paramètres étudiés, et ne se traduisent pas nécessairement par une amélioration franche de l'intelligibilité ou de l'efficacité communicationnelle (Pinto et al., 2005 ; Schulz & Grant, 2000 ; Tykalová et al., 2022). Ainsi, même lorsque l'état moteur général s'améliore, la parole peut rester peu modifiée, voire relativement résistante au traitement. Lorsque la réponse à la lévodopa devient insuffisante ou trop fluctuante, une option chirurgicale peut être envisagée : la stimulation cérébrale profonde (SCP), le plus souvent au niveau du noyau sous-thalamique. Cette intervention peut améliorer certains symptômes moteurs et réduire les fluctuations, mais ses effets sur la parole sont notoirement hétérogènes. Chez certains patients, les troubles de la parole restent stables, ou s'améliorent de manière limitée ; chez d'autres, on observe au contraire une baisse de l'intelligibilité, l'apparition ou l'aggravation d'une dysarthrie, ou encore une dégradation de la qualité prosodique de la parole (Baudouin et al., 2023 ; Pinto et al., 2005, 2023 ; Skodda, 2012 ; Tripoliti et al., 2011). La SCP ne peut donc pas être considérée comme une solution fiable aux troubles de la parole dans la MP.

Ce point est essentiel pour la présente thèse. Il signifie que les troubles de la parole ne sont ni secondairement corrigés de façon suffisante par les traitements dopaminergiques, ni systématiquement améliorés par les approches neurochirurgicales. Dès lors, leur prise en soin repose largement sur des interventions non médicales, et en premier lieu sur la rééducation orthophonique, qui cible directement la parole et le langage, ses effets fonctionnels et ses retentissements communicationnels.

En ce qui concerne la prise en soin orthophonique dans la MP, celle-ci ne vise pas seulement à corriger des symptômes isolés, mesurés en situation d'exercice. Elle poursuit des objectifs fonctionnels plus larges : améliorer l'intelligibilité, réduire l'effort nécessaire pour parler et pour être compris, soutenir la participation dans les échanges du quotidien, et favoriser le transfert des acquis vers des situations écologiques. L'objectif dépasse donc la performance obtenue en séance : il s'agit de favoriser une parole utilisable dans des situations ordinaires, où les contraintes interactionnelles, la fatigue, l'environnement sonore et les

ressources attentionnelles peuvent varier fortement. En France, le guide de parcours de soins de la HAS souligne l'intérêt d'une prise en soin précoce, intensive, pluridisciplinaire et adaptée à l'évolution de la maladie, intégrant le patient et, lorsque cela est pertinent, ses proches (Haute Autorité de Santé, 2016). Plus récemment, un consensus international sur la rééducation dans la MP a également rappelé que les interventions rééducatives devraient faire partie intégrante de la prise en charge tout au long de la maladie, depuis le diagnostic jusqu'aux stades avancés, dans une logique centrée sur la personne, les besoins fonctionnels et les décisions partagées (Goldman et al., 2024).

Les synthèses de la littérature montrent que les approches comportementales peuvent produire des effets positifs sur plusieurs dimensions de la parole, avec des résultats particulièrement robustes pour l'intensité vocale, mais également des effets sur l'intelligibilité et, plus inégalement, sur certains paramètres prosodiques ou articulatoires (Atkinson-Clement et al., 2015 ; Muñoz-Vigueras et al., 2021). Ces résultats soulignent ainsi que la prise en soin orthophonique occupe une place centrale dans la prise en charge lorsque la cible thérapeutique concerne directement la parole.

Parmi les interventions spécifiquement développées pour la dysarthrie hypokinétique, la méthode LSVT LOUD reste la mieux documentée et la plus fréquemment considérée comme référence clinique (Ramig et al., 2001). Il s'agit d'une approche intensive, centrée en première intention sur l'hypophonie, qui combine des séances rapprochées avec un thérapeute et un entraînement quotidien à domicile. Son principe est de pousser le patient à produire une voix plus forte que son niveau habituel, à partir d'une consigne simple et explicite. L'objectif n'est pas uniquement de produire davantage d'intensité en séance, mais de modifier durablement l'échelle interne de calibration de l'effort vocal. Cette logique s'appuie sur l'idée que, dans la MP, le problème ne relève pas seulement d'une capacité motrice réduite, mais aussi d'une difficulté à évaluer correctement le niveau d'effort nécessaire pour atteindre une cible vocale fonctionnelle. Plusieurs travaux ont montré que les personnes vivant avec la MP peuvent sous-estimer l'écart entre leur production réelle et le volume requis pour être entendues efficacement, ou encore juger comme « normal » un niveau d'intensité objectivement diminué (Ho et al., 1999, 2000). Dans ce cadre, l'intensification de l'effort et l'usage de consignes explicites ne valent pas seulement comme techniques de renforcement ; elles visent aussi à recalibrer une boucle sensorimotrice devenue moins fiable (Fox et al., 2006). Les

bénéfices de LSVT LOUD sur l'intensité vocale sont bien établis. Des études de suivi ont montré que les gains obtenus peuvent se maintenir à moyen terme, voire au-delà, notamment dans les formes légères à modérées de la maladie (Ramig et al., 2001 ; Pu et al., 2021). Il importe toutefois de préciser que, même si la littérature sur l'intensité est la plus solide, les effets du traitement ne se limitent pas nécessairement à cette seule dimension. Des améliorations ont également été rapportées sur l'intelligibilité, certaines dimensions articulatoires et la qualité vocale (Levy et al., 2020 ; Moya-Galé et al., 2024 ; Sapir et al., 2007). Ces effets ne sont pas toujours de même force selon les études, mais ils montrent qu'une intervention centrée initialement sur l'intensité peut avoir des répercussions plus larges sur l'organisation de la parole.

C'est précisément là qu'apparaît une limite importante de la prise en soin actuelle. Si LSVT LOUD peut avoir des effets indirects sur certains paramètres suprasegmentaux, la prosodie n'en constitue pas l'objectif principal. Dans les revues consacrées aux traitements comportementaux de la parole dans la MP, les effets les mieux documentés concernent avant tout l'intensité vocale, avec des retombées plus variables sur d'autres dimensions de la parole (Atkinson-Clement et al., 2015 ; Martens et al., 2015). En parallèle, les enquêtes menées auprès d'orthophonistes montrent que la prosodie demeure relativement peu évaluée et peu travaillée de manière spécifique, y compris lorsqu'un trouble prosodique est suspecté (Hawthorne & Fischer, 2020 ; Petitpain et al., 2024). Cette situation contraste avec le rôle fonctionnel majeur des dimensions prosodiques. La place encore limitée accordée à la prosodie dans les approches thérapeutiques apparaît ainsi comme un enjeu important pour la prise en soin des troubles de la parole dans la MP.

Dans nos travaux, l'enjeu n'est pas d'opposer une orthophonie centrée sur la parole et la voix à une approche centrée sur la prosodie. Il s'agit plutôt de montrer que certaines limites des approches actuelles ouvrent un espace pour d'autres cibles thérapeutiques. L'hypothèse d'une altération de la calibration sensorielle, souvent mobilisée pour expliquer les difficultés de régulation de l'intensité, peut aussi être discutée pour des dimensions plus larges de la parole, notamment certaines composantes fréquentielles et temporelles qui structurent la prosodie : variabilité de F0, débit, durées syllabiques ou pauses. À cet égard, les modèles récents de la production de la parole invitent à penser ensemble contrôle moteur, feedback

sensoriel et organisation temporelle de la parole, plutôt qu'à isoler une seule variable acoustique (Manes et al., 2024).

C'est dans cette perspective que l'utilisation d'indices rythmiques externes devient particulièrement intéressante. Sans se substituer aux approches orthophoniques validées, elle permet de repenser autrement la question de la prise en soin des dimensions prosodiques de la parole, et notamment rythmiques. Elle conduit aussi à envisager la rééducation non seulement comme une correction d'un symptôme acoustique, mais également comme une tentative de soutien aux mécanismes de structuration temporelle et sensorimotrice impliqués dans la production de la parole. C'est dans ce cadre, entre les interrogations sur les mécanismes sous-jacents, les apports établis de l'orthophonie et les limites des traitements actuellement disponibles pour la prosodie, que s'inscrit le présent travail.

Chapitre 2
Rythme, interventions et amorçage rythmique

2.1 Interventions basées sur le rythme dans la maladie de Parkinson

Dans la MP, les difficultés motrices ne sauraient être réduites à une simple diminution de force ou d'amplitude. Elles concernent également l'organisation temporelle de l'action, c'est-à-dire la capacité à initier une séquence, à en maintenir la régularité et à en ajuster finement le déroulement au cours de l'exécution (Nombela et al., 2013 ; Puyjarinet et al., 2019 ; Tinaz et al., 2016). Cette dimension temporelle est engagée dans de nombreuses activités de la vie quotidienne, telles que la marche, les mouvements répétés ou, plus largement, toute situation requérant une coordination stable dans le temps.

Dans cette perspective, plusieurs auteurs ont proposé d'interpréter une partie des manifestations motrices de la MP comme l'expression d'un trouble plus général du timing, affectant la production et la stabilisation de régularités temporelles internes (Nombela et al., 2013 ; Puyjarinet et al., 2019). Un point particulièrement important, du point de vue clinique, est que cette fragilité du timing interne ne s'accompagne pas nécessairement d'une altération équivalente de l'utilisation des repères temporels externes. Au contraire, chez les personnes atteintes de la MP, la capacité à s'ajuster à un signal rythmique externe (qu'il s'agisse d'un métronome, d'un rythme musical ou d'un autre indice auditif périodique) peut demeurer au moins partiellement préservée, en particulier dans certaines tâches motrices répétitives ou guidées par des indices auditifs (Lim et al., 2005 ; Koshimori & Thaut, 2018). Cette association entre difficulté de génération interne et meilleure utilisation de repères externes constitue l'un des principaux arguments en faveur des approches fondées sur le rythme. Leur intérêt réside moins dans la restauration d'un timing interne déficitaire que dans la mise à disposition d'un cadre temporel externe, susceptible de soutenir ou stabiliser l'organisation de l'action (Nombela et al., 2013 ; Puyjarinet et al., 2019 ; Koshimori & Thaut, 2018).

La stimulation auditive rythmique (SAR) illustre particulièrement bien ce principe (Thaut & Abiru, 2010). Dans ce type d'approche, un signal rythmique régulier, simple ou musical, est utilisé comme repère temporel pour soutenir l'exécution motrice et favoriser l'alignement du mouvement sur une trame externe. Le rythme n'est alors pas conçu comme un simple accompagnement, mais comme une structure susceptible de réduire la variabilité temporelle, de faciliter l'initiation et de soutenir la coordination séquentielle du geste (Thaut et al., 1996 ; McIntosh et al., 1997). Les synthèses disponibles décrivent d'ailleurs la SAR comme l'une des approches les mieux établies dans le champ des interventions musicales et rythmiques en MP, en particulier pour la marche et l'équilibre (Lim et al., 2005 ; Koshimori & Thaut, 2023 ; Nieuwboer et al., 2007). C'est en effet dans le domaine locomoteur que cette logique a été la plus largement documentée. Des études ont montré que la stimulation auditive rythmique pouvait améliorer des paramètres fonctionnels tels que la vitesse de marche, la cadence, la longueur du pas ou la symétrie, y compris chez des patients testés dans différents états médicamenteux (Thaut et al., 1996 ; McIntosh et al., 1997). Certaines données suggèrent également qu'un tempo légèrement supérieur à la cadence spontanée peut produire des effets plus marqués chez certains patients, et que les bénéfices observés peuvent persister après l'arrêt immédiat du signal, avant de diminuer progressivement en l'absence de poursuite de l'entraînement (McIntosh et al., 1997). Plus largement, les revues disponibles concluent à un bénéfice particulièrement robuste des indices auditifs sur la vitesse de marche, même si la généralisation à d'autres paramètres ou à la vie quotidienne dépend des protocoles et reste à discuter (Lim et al., 2005 ; Nombela et al., 2013 ; Ye et al., 2022). Cette prudence est d'autant plus nécessaire que les compétences rythmiques ne constituent pas un ensemble homogène, mais renvoient à des composantes perceptives et sensorimotrices partiellement dissociables, dont l'évaluation peut être affinée à l'aide d'outils dédiés comme la BAASTA (Dalla Bella et al., 2024). Il convient toutefois de noter qu'un entraînement régulier fondé sur des indices rythmiques peut entraîner des bénéfices persistant plusieurs semaines après l'intervention, même si leur maintien dépend des protocoles et des paramètres considérés (Nieuwboer et al. 2007) et que les capacités rythmiques sont également améliorées (Dalla Bella et al., 2017a ; Benoit et al., 2014 ; Bourdon et al., 2025*)*. Cette approche a également été prolongée par le développement d'outils numériques utilisables en contexte plus écologique, tels que BeatWalk, puis BeatMove, développés et évalués principalement chez des personnes atteintes de la MP afin de proposer une stimulation

musicale personnalisée et synchronisée en temps réel à la marche (Cochen De Cock et al., 2021 ; Bourdon et al., 2025).

D'autres dispositifs fondés sur le rythme ont également été proposées, au-delà de la SAR, afin de cibler plus directement les compétences rythmiques altérées dans la MP. C'est le cas de Rhythm Workers, un jeu sérieux musical conçu pour entraîner les capacités rythmiques (Bégel et al., 2018). Chez des personnes atteintes de la MP, un entraînement à domicile avec ce type d'outil a montré des effets bénéfiques sur les performances rythmiques, y compris dans des tâches manuelles et oro-faciales, suggérant que l'effet de l'entraînement pourrait dépasser le seul domaine de la marche et concerner plus largement certaines composantes du contrôle moteur séquentiel (Puyjarinet et al., 2022). Dans un registre différent, des approches fondées sur la sonification du geste ont également été proposées pour la prise en soin de la dysgraphie et de la micrographie dans la MP, notamment à travers la sonification musicale du mouvement d'écriture (Véron-Delor et al., 2020 ; Véron-Delor, 2021). Enfin, au-delà des usages compensatoires ou rééducatifs, l'évaluation des compétences rythmiques des personnes atteintes de la MP occupe une place importante dans la littérature scientifique et devient plus accessible à la clinique grâce à des outils dédiés tels que la BAASTA, développée pour évaluer les capacités perceptives et sensorimotrices de timing, puis adaptée plus récemment sous forme mobile (Dalla Bella et al., 2017b, 2024).

Ainsi, les interventions fondées sur le rythme disposent aujourd'hui d'un socle empirique particulièrement solide dans le champ moteur, en particulier locomoteur. Cette efficacité explique la place qu'elles occupent dans la réflexion thérapeutique dans la MP. Elle ne suffit toutefois pas à garantir un bénéfice comparable dans d'autres domaines moteurs.

2.1.1. Cadre neurofonctionnel : ganglions de la base et cervelet

Les effets des interventions rythmiques observés en clinique dans la MP invitent également à préciser les mécanismes neurofonctionnels susceptibles de les sous-tendre. Dans cette perspective, plusieurs travaux ont mis en avant le rôle complémentaire des ganglions de la base et du cervelet dans le contrôle temporel de l'action. Les ganglions de la base sont généralement associés à l'initiation de l'action, à la sélection des programmes moteurs et à certains aspects du timing interne, tandis que le cervelet est davantage impliqué dans l'ajustement temporel fin, la coordination, le traitement de l'erreur et l'adaptation en ligne (Teki et al., 2012 ; Tseng et al., 2007). Cette distinction ne doit toutefois pas être comprise comme un partage strict et exclusif des fonctions, mais plutôt comme une manière de décrire deux pôles fonctionnels complémentaires au sein de réseaux fortement interconnectés.

Dans le champ du traitement temporel, plusieurs auteurs ont proposé de distinguer deux voies, dont les contributions peuvent varier selon les tâches et les contextes. La première, dite *beat-based*, est généralement associée au réseau cortico-ganglio-basal-thalamo-cortical et serait plus particulièrement impliquée dans le traitement de régularités temporelles organisées autour d'une pulsation, contribuant ainsi à la prédiction temporelle et à la stabilité du séquençage moteur (Teki et al., 2011, 2012 ; Nozaradan et al., 2017 ; Criscuolo et al., 2025). La seconde, dite *duration-based*, plus étroitement liée au cervelet, contribuerait davantage à la calibration des intervalles et à la précision de l'exécution (Teki et al., 2011, 2012 ; Breska & Ivry, 2018 ; Criscuolo et al., 2025). Dans cette perspective, un rythme régulier peut agir comme un repère temporel externe, susceptible de soutenir soit la stabilité du séquençage, soit la réduction de la variabilité temporelle au cours de l'exécution.

Dans la MP, où l'atteinte du système ganglio-basal est susceptible de fragiliser certains mécanismes d'initiation, de séquençage et d'organisation temporelle interne (McGregor & Nelson, 2019), ce cadre permet de mieux comprendre pourquoi des repères rythmiques externes peuvent soutenir l'action sans pour autant relever d'un mécanisme simple et univoque. Les bénéfices observés ont ainsi été interprétés soit comme le reflet d'un renforcement compensatoire de l'implication du circuit cérébello-thalamo-cortical, soit comme l'expression d'une exploitation plus efficace de l'activité résiduelle des boucles ganglio-thalamo-corticales (Koshimori & Thaut, 2018 ; Nombela et al., 2013). Il ne s'agit pas

de postuler un simple mécanisme de substitution, mais plutôt d'envisager que le rythme externe reconfigure les conditions de fonctionnement du système en fournissant une information temporelle prédictible susceptible de stabiliser la dynamique de l'action.

Ce cadre neurofonctionnel est important pour la présente thèse, car il fournit un point d'appui théorique pour examiner dans quelles conditions une structure rythmique perçue peut influencer la préparation et l'organisation de comportements moteurs complexes. Les chapitres 2 et 3 de la deuxième partie mobilisent ce cadre pour étudier les effets de l'amorçage rythmique sur le temps d'initiation, le débit et certaines dimensions prosodiques de la parole, en comparant des personnes atteintes de la MP, des personnes atteintes d'ataxie spinocérébelleuse (ASC) et des participants contrôles.

2.1.2. Transfert à la parole

L'extension des approches rythmiques à la parole repose sur le fait que la parole constitue, elle aussi, une activité motrice fortement organisée dans le temps. Sa production suppose une coordination fine entre respiration, phonation et articulation, ainsi qu'une structuration séquentielle suffisamment stable pour garantir l'intelligibilité, le caractère naturel de la parole et, plus largement, l'efficacité de la communication. Dans cette perspective, il est envisageable qu'un repère temporel externe puisse soutenir non seulement la marche ou d'autres comportements moteurs, mais aussi certaines dimensions de la parole. Cette transposition ne va toutefois pas de soi. En effet, si des questions de structuration temporelle et d'adaptation au contexte se posent également dans la locomotion, la parole mobilise en outre une hiérarchie prosodique, des contraintes linguistiques et un ajustement interactionnel qui compliquent toute transposition directe.

Les données disponibles dans ce domaine restent nettement moins nombreuses que celles portant sur la marche, mais elles suggèrent néanmoins qu'un guidage rythmique peut modifier certaines dimensions de la parole dans la MP. Dans une étude pionnière, Thaut et al. (2001) rapportent ainsi que la présentation d'indices auditifs rythmiques peut influencer le contrôle moteur de la parole chez des personnes atteintes de la MP. Ce travail s'inscrit dans le prolongement des approches cliniques fondées sur le rythme et développées pour le

mouvement, en montrant que la même logique peut aussi être envisagée dans le champ de la parole. Même s'il ne permet pas, à lui seul, de préciser les mécanismes en jeu ni l'ampleur clinique des effets observés, il constitue un premier point d'appui important pour penser le transfert des approches rythmiques vers la production verbale.

Les travaux utilisant explicitement le rythme dans le domaine de la parole sont donc peu nombreux. Il existe néanmoins un certain nombre d'études ayant utilisé des stimuli verbaux plutôt que musicaux ou rythmiques. Cette caractéristique est importante, car elle signifie que les effets observés ne peuvent pas être uniquement attribués à une structuration temporelle abstraite : dans plusieurs paradigmes, l'amorce fournit aussi des indices prosodiques, segmentaux et interactionnels. La littérature disponible suggère que dans la MP, une structure temporelle préalable peut influencer la mise en route de la production et certaines formes d'accommodation rythmique à la parole d'autrui. L'étude de Späth et al. (2016) constitue, à cet égard, un point d'appui important. Dans ce paradigme, des participants présentant une dysarthrie hypokinétique liée à la MP et des participants contrôles entendaient une « phrase-amorce » produite par un locuteur servant de modèle, puis lisaient à voix haute une phrase-cible, régulière ou irrégulière sur le plan métrique. Les auteurs montrent que les amorces métriquement régulières conduisent à des initiations de parole plus précoces que les amorces irrégulières, autrement dit à des temps d'initiation plus courts par rapport à la fin de l'énoncé entendu. Ils observent également une accommodation rythmique plus marquée entre la phrase produite et la phrase-amorce lorsque les cibles présentent elles-mêmes une structure métrique régulière. Fait particulièrement intéressant pour la présente thèse, cet effet apparaît chez les personnes atteintes de la MP, ce qui suggère qu'elles peuvent exploiter des indices métriques externes présents dans la parole de l'interlocuteur pour ajuster leur propre production.

Cette lecture est renforcée par les résultats d'Aichert et al. (2021). Dans une tâche de complétion de phrase, des participants présentant une dysarthrie hypokinétique, une apraxie de la parole ou un trouble phonologique devaient poursuivre la production d'un modèle auditif en maintenant son rythme et sa fluidité. Les auteurs observent que, chez les personnes atteintes de la MP, les temps d'initiation de la parole et la sensibilité aux régularités rythmiques du modèle restent globalement proches de ceux des témoins, et que les patients peuvent tirer profit de la régularité du premier stimulus. Les auteurs concluent ainsi que,

malgré les troubles moteurs de la parole, les personnes atteintes de la MP demeurent capables d'anticiper temporellement et de s'accommoder à un patron rythmique régulier entendu. Il convient toutefois de noter que ce travail ne relève pas d'un amorçage musical au sens strict. Il repose plutôt sur une exposition à un modèle auditif dont les participants doivent prolonger la structure temporelle. L'effet observé renvoie donc davantage à une accommodation au rythme d'un modèle entendu qu'à un effet d'amorçage rythmique musical sur une production de parole autonome. Dans le même sens, Späth et al. (2022), dans un paradigme d'adaptation au débit et au rythme de la parole d'autrui, montrent que les participants atteints de la MP conservent, et parfois accentuent, leur propension à s'ajuster au débit et au rythme d'un modèle, alors que cette capacité apparaît largement réduite dans la dégénérescence cérébelleuse. Cette étude suggère elle aussi que la sensibilité aux structures temporelles externes en parole reste mobilisable dans la MP. En parallèle, d'autres travaux laissent penser que le rythme peut également agir au-delà de la seule exécution motrice. Kotz et Gunter (2015) ont ainsi montré, dans une étude à cas unique, qu'un indice auditif métriquement aligné avec la structure accentuelle de la phrase pouvait améliorer le traitement syntaxique et sémantique d'énoncés. Il s'agit toutefois d'un travail centré sur la compréhension et non sur la production.

Dans la MP, les données disponibles restent donc encore limitées, les paradigmes sont hétérogènes, les amorces majoritairement verbales, et les mesures ne portent pas toujours sur les mêmes dimensions de la parole. Surtout, les effets sur des marqueurs explicitement prosodiques demeurent encore peu documentés. Dans ce contexte, il paraît particulièrement pertinent d'examiner plus précisément le paradigme d'amorçage. En dissociant temporellement l'exposition au rythme et la production de la parole, ce paradigme offre un cadre adapté pour étudier si une structure temporelle perçue en amont peut ensuite influencer la parole. C'est ce paradigme, ses fondements théoriques, ses principales variantes et les effets qui lui sont attribués, qui seront maintenant examinés plus en détail dans la section suivante.

2.2. L’amorçage rythmique

L’amorçage rythmique désigne un paradigme dans lequel l’exposition à un stimulus rythmiquement structuré précède une tâche de langage ou de parole, avec l’hypothèse que cette exposition peut influencer la performance ultérieure. À la différence des approches fondées sur un repère temporel présenté pendant l’action, l’amorçage ne vise pas d’abord une synchronisation en temps réel de l’exécution. Il s’agit plutôt d’agir en amont de la tâche, en modifiant les conditions initiales du traitement par la préparation temporelle, la prédictibilité du contexte auditif et, plus largement, la préparation sensorimotrice, de façon à ce qu’un effet puisse persister au-delà de la stimulation elle-même. Cette logique est au cœur de plusieurs travaux comme ceux par exemple de Cason et Schön (2012), de Przybylski et al. (2013) ou encore de Cason et al. (2015a), qui montrent qu’une exposition préalable à une structure temporelle peut moduler des performances subséquentes de traitement verbal. Plus généralement, ce cadrage repose sur l’idée que la parole, comme la musique, est un signal qui se déploie dans le temps et dont le traitement dépend en partie de la capacité du système à extraire des régularités, à segmenter des groupements pertinents et à former des attentes sur les événements à venir. Une amorce rythmique ne constitue donc pas un simple contexte sonore précédant la cible : elle peut fournir une structure temporelle susceptible d’orienter l’attention, de soutenir les mécanismes de prédiction et de faciliter ensuite le traitement d’un stimulus linguistique ou la préparation d’une production verbale. Cette lecture est basée sur la *Dynamic Attending Theory* (DAT), selon laquelle des oscillations attentionnelles peuvent s’aligner sur la structure temporelle d’événements externes et renforcer le traitement aux moments attendus (Large & Jones, 1999). Elle s’accorde aussi avec des travaux soulignant le rôle des oscillations neuronales et des régularités rythmiques de la parole dans la compréhension (Peelle & Davis, 2012). Surtout, plusieurs résultats suggèrent que cet alignement ne s’interrompt pas nécessairement avec la fin du stimulus : il peut persister sur quelques cycles et ainsi influencer le traitement perceptif du matériel linguistique qui suit, dans une logique d’effet d’écho ou d’entraînement soutenu (Kösem et al., 2018 ; Zoefel, 2018).

Dans les travaux consacrés à l’amorçage rythmique, l’effet du rythme est le plus souvent étudié à partir d’une comparaison entre des amorces régulières et des amorces irrégulières, parfois complétée par une condition contrôle silencieuse ou non musicale

(Przybylski et al., 2013 ; Canette et al., 2019 ; Fiveash et al., 2020). Cette opposition permet d'évaluer dans quelle mesure un contexte temporel prédictible peut modifier le traitement d'un stimulus linguistique présenté ensuite, notamment dans des tâches de perception phonologique, de jugement grammatical ou de compréhension de la parole (Cason & Schön, 2012 ; Przybylski et al., 2013 ; Berthault et al., 2024). Elle doit toutefois être interprétée avec prudence, car les termes de régularité et d'irrégularité ne renvoient pas toujours aux mêmes propriétés acoustiques ou métriques selon les études. Une amorce régulière peut ainsi reposer sur une pulsation isochrone, sur une organisation métrique hiérarchisée, sur une saillance accentuelle récurrente ou simplement sur une stabilité temporelle plus marquée. À l'inverse, une amorce irrégulière peut correspondre à une variabilité accrue des intervalles, à une perturbation de la périodicité ou à une structure moins facilement anticipable. L'enjeu, dans les travaux portant sur l'amorçage rythmique, n'est donc pas seulement d'opposer des amorces régulières et irrégulières. Il est aussi d'identifier les propriétés effectivement manipulées (régularité des intervalles, organisation métrique, saillance accentuelle, tempo, durée d'exposition ou congruence avec la cible) et de déterminer à quel niveau elles sont susceptibles d'agir : orientation attentionnelle, prédiction temporelle, préparation sensorimotrice ou organisation de la réponse verbale (Cason et al., 2015a ; Falk et al., 2017 ; Fiveash et al., 2021 ; György et al., 2024).

La nature des amorces constitue une autre source importante de variation entre les paradigmes. Certaines études utilisent des séquences musicales ou non verbales, afin d'isoler autant que possible la structure temporelle du stimulus, tandis que d'autres s'appuient sur des stimuli verbaux, qui fournissent simultanément des informations rythmiques, prosodiques, segmentales et parfois interactionnelles (Cason & Schön, 2012 ; Falk & Dalla Bella, 2016 ; Späth et al., 2016). De même, l'amorce peut être uniquement auditive ou s'inscrire dans une situation auditivo-motrice, lorsque l'écoute est associée à une synchronisation gestuelle ou à un entraînement préalable (Berthault et al., 2024 ; Cason et al., 2015a ; Falk & Dalla Bella, 2016 ; Falk et al., 2017 ; Gindre, 2024). Ces distinctions sont importantes, car elles modifient la nature du couplage engagé entre perception, prédiction temporelle et préparation de l'action, dimensions qui occupent une place centrale dans les cadres théoriques récents sur les liens entre rythme musical et traitement de la parole (Large & Jones, 1999 ; Patel & Iversen, 2014 ; Fiveash et al., 2021). Elles le sont d'autant plus dans le cadre de cette thèse, où l'objectif n'est pas d'évaluer un effet général du rythme sur le

langage, mais de déterminer si une structure temporelle entendue avant la parole peut influencer certaines dimensions de la production orale, en particulier dans une pathologie où les mécanismes d'initiation, de séquençage et de contrôle temporel des gestes moteurs sont fragilisés. L'amorçage rythmique est donc envisagé ici comme un outil expérimental permettant de dissocier l'exposition au rythme de l'exécution elle-même, et d'interroger plus finement les conditions dans lesquelles une régularité temporelle préalable peut se transférer et interagir avec les mécanismes de la production de la parole.

Enfin, ces paradigmes diffèrent également par la durée des amorces et par l'échelle à laquelle l'effet est attendu. Dans les paradigmes de type *global rhythmic priming*, développés notamment dans les travaux de Tillmann et collègues, les amorces sont généralement longues (souvent de l'ordre de 30 secondes ou plus) et visent à induire un état de préparation temporelle susceptible d'influencer le traitement d'une série d'items subséquents, en particulier dans des tâches de jugement grammatical ou de traitement syntaxique (Fiveash et al., 2020 ; Przybylski et al., 2013). À l'inverse, d'autres paradigmes reposent sur des amorces plus brèves et sur une relation plus étroite entre une amorce donnée et la cible qui suit, comme dans les travaux sur la perception phonologique, la congruence métrique ou l'alignement prosodique (Cason & Schön, 2012 ; Cason et al., 2015a ; Falk et al., 2017). Cette distinction est d'autant plus importante que, dans les travaux sur le traitement syntaxique, les amorces musicales longues et régulières sont souvent supposées agir non seulement sur la prédictibilité temporelle du signal, mais aussi sur sa structuration métrique hiérarchique. Dans cette perspective, l'effet observé ne tiendrait pas uniquement à une simple régularité temporelle, mais aussi à la capacité de la musique à induire un cadre métrique hiérarchisé susceptible de faciliter, dans certaines tâches, le traitement de structures linguistiques elles-mêmes organisées hiérarchiquement (Fiveash et al., 2020 ; György et al., 2024 ; Przybylski et al., 2013).

L'amorçage rythmique apparaît ainsi moins comme une procédure unitaire que comme une famille de paradigmes, dont les effets dépendent notamment de la nature des amorces, de leur durée, de leur organisation temporelle et de leur articulation avec la tâche langagière visée (Fiveash et al., 2025).

La question de la congruence mérite d'ailleurs d'être distinguée de celle de la régularité. Une amorce peut être globalement régulière sans pour autant préparer au mieux la cible si sa structure métrique ou prosodique ne correspond pas à celle qui sera ensuite traitée. Inversement, plusieurs études montrent qu'une proximité structurelle entre amorce et cible renforce le bénéfice observé, qu'il s'agisse de perception phonologique, de traitement verbal ou de compréhension de la parole dans le bruit, comme l'illustrent Cason et Schön (2012) et Cason et al. (2015a). Cette distinction est centrale pour cette thèse, car elle conduit à ne pas penser l'effet du rythme uniquement comme une simple « mise en cadence » globale, mais aussi comme un effet possible de préparation plus spécifique de certaines régularités attendues. Au-delà de ces distinctions, l'amorçage rythmique présente un intérêt méthodologique et translationnel particulier. Parce que la stimulation précède la tâche, il permet de mieux distinguer un éventuel effet de préparation temporelle d'un simple effet d'indiçage concomitant. Cette propriété est particulièrement pertinente lorsqu'on s'intéresse à la parole, et plus encore à la parole dans la MP, car elle permet de tester si une simple exposition préalable à une structure rythmique peut influencer la production ou le traitement sans imposer un signal pendant l'exécution. Dans cette perspective, l'amorçage rythmique peut être envisagé comme un format expérimental intermédiaire entre les recherches fondamentales sur la prédiction temporelle et les approches cliniques fondées sur le rythme. C'est précisément ce qui en fait un cadre particulièrement utile pour les travaux présentés dans ce manuscrit.

2.2.1. Effets sur la perception et le traitement linguistique

Les travaux ayant mis en évidence un effet de l'amorçage rythmique sur le langage ont principalement concerné la perception et, plus largement, des tâches de traitement linguistique réalisées après une exposition préalable à une structure temporelle. Dans ce cadre, l'hypothèse générale est qu'une amorce rythmiquement organisée peut préparer le système à traiter plus efficacement certains événements langagiers subséquents, en particulier lorsque la cible présente elle-même une organisation temporelle compatible avec celle de l'amorce.

Sur le versant phonologique, Cason et Schön (2012) ont montré qu'une amorce musicale simple pouvait améliorer le traitement de pseudo-mots parlés. Leur étude, combinant mesures comportementales et électrophysiologiques, suggère que les attentes temporelles générées par l'amorce influencent le traitement du mot qui suit, avec des effets observés sur plusieurs composantes ERP (potentiel évoqué) ainsi que sur les temps de réponse. Les auteurs interprètent ces résultats comme la mise en évidence d'un effet croisé entre rythme musical et rythme de la parole, l'amorce facilitant le traitement phonologique lorsqu'elle est compatible avec certaines propriétés prosodiques du stimulus verbal. Dans la même logique, Cason et al. (2015a) ont montré que des amorces rythmiques et des entraînements auditivo-moteurs pouvaient également modifier la perception de la parole, ce qui renforce l'idée qu'une structuration temporelle préalable peut façonner le traitement perceptif du signal verbal.

Les effets décrits ne se limitent toutefois pas au niveau phonologique. Plusieurs travaux suggèrent qu'un amorçage rythmique peut aussi influencer des traitements linguistiques de plus haut niveau, notamment en syntaxe (p. ex. Canette et al., 2019 ; Fiveash et al. 2020 ; Przybylski et al., 2013). Przybylski et al. (2013) ont ainsi observé, chez des enfants présentant un trouble développemental du langage ou une dyslexie, une amélioration des performances à une tâche de jugement de grammaticalité après une amorce musicale régulière, comparativement à une amorce irrégulière. L'intérêt de ce résultat n'est pas seulement de mettre en évidence un bénéfice ponctuel, mais de suggérer qu'une structure temporelle préalable peut soutenir des traitements linguistiques qui mobilisent une organisation hiérarchique des unités, comme c'est le cas en syntaxe. En effet, le traitement syntaxique ne repose pas uniquement sur l'identification successive de mots isolés, mais sur leur intégration dans des constituants de niveaux différents, eux-mêmes organisés au sein d'une structure phrastique plus large. Les mots sont ainsi regroupés en unités intermédiaires, telles que des groupes nominaux ou verbaux, dont l'interprétation dépend à la fois de leur position dans la phrase et des relations qu'ils entretiennent avec les autres constituants. Dans cette perspective, l'effet d'un amorçage rythmique sur le jugement de grammaticalité peut être interprété comme le signe que les régularités temporelles ne facilitent pas seulement le traitement local des sons ou des syllabes, mais peuvent également soutenir la mise en place de cadres d'anticipation utiles à l'intégration d'unités linguistiques plus larges.

Dans une perspective proche, Kotz et Gunter (2015) rapportent, dans le contexte de la MP, qu'un indice auditif métriquement aligné avec la structure accentuelle de la parole peut améliorer le traitement syntaxique et sémantique. Même si cette étude repose sur un design particulier et sur un nombre de participants limité, elle soutient l'idée que le rythme peut interagir avec des niveaux de traitement linguistique variés, y compris lorsque ceux-ci impliquent l'intégration de relations structurales entre éléments linguistiques. Plus largement, ces résultats invitent à considérer le rythme comme un support potentiel de structuration temporelle du traitement langagier : il ne déterminerait pas directement l'analyse syntaxique, mais pourrait favoriser l'allocation de l'attention dans le temps et faciliter l'intégration progressive des unités linguistiques au sein d'une organisation hiérarchique.

Ces données sont importantes pour cette thèse, car elles montrent que l'amorçage rythmique peut influencer des traitements perceptifs et linguistiques variés, du phonologique au syntaxique, sans qu'un guidage temporel soit maintenu pendant la tâche. Elles suggèrent aussi que les effets observés ne concernent pas un phénomène unique : selon les paradigmes, ils peuvent refléter des modifications de l'attention temporelle, des attentes métriques, de la préparation sensorimotrice ou de l'intégration de la structure temporelle au traitement linguistique. Cette diversité explique à la fois l'intérêt du paradigme et la prudence nécessaire dans son interprétation. Elle justifie également de s'interroger, dans un second temps, sur l'extension de ces effets à la production de la parole elle-même.

2.2.2. Extension à la production

Au-delà du versant perceptif, plusieurs travaux suggèrent que l'amorçage rythmique peut également influencer la production. Cette extension est importante pour comprendre si, et dans quelle mesure, l'amorçage rythmique peut modifier la planification, l'initiation ou certains paramètres de l'exécution de la parole. Dans ce champ, la littérature reste moins homogène que dans le domaine de la perception, tant du point de vue des tâches employées que des dimensions analysées.

Une première série de travaux suggère que l'amorçage rythmique peut affecter la précision de la production. Chez des enfants sourds équipés d'aides auditives, Cason et al.

(2015b) ont observé qu'une amorce musicale métriquement congruente avec la phrase cible améliorait l'exactitude phonologique de la répétition, en comparaison avec des conditions de base ou d'incongruence. Dans une autre population clinique, et avec des stimuli plus longs, Fiveash et al. (2023) ont montré que des amorces musicales régulières pouvaient également améliorer la répétition de phrases chez des enfants présentant un trouble développemental du langage, avec un bénéfice observé sur l'exactitude grammaticale. Ces résultats suggèrent qu'une structure temporelle préalable peut soutenir non seulement la précision phonologique, mais aussi, dans certaines tâches de répétition, des dimensions plus larges de l'organisation linguistique de la réponse.

D'autres travaux mettent davantage en évidence des effets sur les paramètres temporels de la production, en particulier la latence d'initiation et le débit. Dans une étude en électroencéphalographie (EEG) en mandarin, Zhang et Zhang (2019) montrent qu'une amorce rythmique peut faciliter la production lorsque le patron rythmique de l'amorce est congruent avec celui de la cible, avec un temps d'initiation plus court dans les conditions congruentes, au moins pour certains patrons rythmiques. Plus récemment, Gindre (2024) suggère que cet effet sur l'initiation de la parole peut être renforcé lorsque l'écoute de l'amorce s'accompagne d'un engagement moteur. Jungers et Hupp (2018) ont montré, chez des adultes comme chez des enfants, que des amorces musicales rapides conduisaient à une production verbale plus rapide que des amorces plus lentes. Enfin, quelques travaux suggèrent que l'amorçage rythmique peut également influencer certaines propriétés globales du rythme produit. Liu et Liu (2024) ont ainsi montré, chez des apprenants chinois du français, que des amorces rythmiques pouvaient rendre la production plus proche du rythme de la langue cible. Ce résultat reste difficilement comparable aux études centrées sur l'initiation ou le débit, mais il suggère que le transfert perception-production ne se limite pas à une simple accélération de la réponse et peut aussi concerner l'organisation temporelle plus globale de la parole produite.

Pris ensemble, ces travaux indiquent donc que l'amorçage rythmique peut influencer la production verbale. Les effets actuellement les mieux documentés concernent l'initiation de la parole, certains paramètres temporels globaux comme le débit, ainsi que la précision de la production, qu'elle soit phonologique ou, dans certains paradigmes de répétition,

grammaticale. Quelques travaux suggèrent également un effet sur le rythme produit lui-même. En revanche, la généralisation à des dimensions prosodiques plus larges demeure, à ce stade, moins bien établie.

L'intérêt de l'amorçage rythmique pour cette thèse tient à la possibilité de tester, sans guidage temporel pendant l'exécution elle-même, si une structure temporelle préalable peut modifier certaines étapes de la production : la préparation de la réponse, son initiation, son organisation temporelle globale ou certains paramètres prosodiques. Cette approche est particulièrement pertinente dans la MP, où l'atteinte des circuits ganglio-basaux peut fragiliser certains mécanismes impliqués dans la mise en route, le séquençage et la régulation temporelle des actions motrices, y compris de la parole, sans que toutes les dimensions de la production orale soient pour autant affectées de manière homogène.

Plusieurs questions demeurent ainsi ouvertes. La littérature ne permet pas encore de préciser quelles composantes de la parole sont les plus sensibles à l'amorçage rythmique, ni comment cet effet varie selon les profils neurologiques. Cette question est particulièrement importante dans la MP, où les données demeurent limitées, souvent hétérogènes, et encore peu informatives sur les dimensions explicitement prosodiques. L'étude de l'amorçage rythmique dans cette pathologie présente donc un double intérêt : elle permet d'évaluer plus directement l'effet d'une structure rythmique préalable sur la parole, et elle offre, lorsqu'elle est mise en regard d'une atteinte cérébelleuse, un moyen d'interroger les mécanismes neurocognitifs susceptibles de sous-tendre cet effet, étant donné que, comme décrit dans la section 2.1.1, ces deux structures pourraient jouer un rôle.

Chapitre 3
Modèles et cadre interprétatif

Pour interpréter les effets de l'amorçage rythmique sur la parole et développer des applications cliniques fondées sur les preuves, il faut préciser dans quel cadre théorique ces effets peuvent être compris, et à quel niveau ils sont supposés intervenir. Cette question est d'autant plus importante que les travaux présentés jusqu'ici mobilisent plusieurs niveaux d'analyse à la fois : la structuration temporelle du signal, la formation de prédictions, les relations entre perception et action, ainsi que la possibilité d'un transfert entre rythme perçu et parole produite. Les effets observés ne prennent donc pleinement sens que s'ils sont replacés dans un ensemble de modèles permettant d'articuler ces différentes dimensions.

Plusieurs cadres issus de la littérature peuvent être mobilisés à cette fin. Certains portent plus largement sur les relations entre musique et langage, comme l'hypothèse OPERA (*Overlap, Precision, Emotion, Repetition, Attention*) de Patel (2011, 2012), selon laquelle l'entrainement musical peut favoriser certaines fonctions impliquées dans la parole en raison du chevauchement partiel des réseaux mobilisés, de l'exigence de précision du traitement musical, ainsi que du rôle de la répétition, de l'attention et de l'émotion. D'autres mettent plus directement l'accent sur les mécanismes temporels et rythmiques, comme l'hypothèse SEP de Fujii et Wan (2014), qui insiste sur le traitement de l'enveloppe sonore (*Sound Envelope Processing)* et sur la synchronisation à une pulsation (*Synchronization and Entrainment to Pulse*), ou encore la DAT de Large et Jones (1999), qui formalise l'idée d'un alignement des dynamiques attentionnelles internes sur la structure temporelle d'un signal externe. À cela s'ajoutent des propositions insistant davantage sur la participation des systèmes moteurs à la perception du rythme, comme l'hypothèse ASAP (*Action Simulation for Auditory Prediction*) de Patel et Iversen (2014).

Ces cadres sont importants, car ils ont contribué à déplacer la réflexion d'une simple analogie entre musique et langage vers une interrogation plus précise sur les mécanismes susceptibles de soutenir un effet du rythme sur la parole. Ils restent toutefois relativement généraux au regard de la question plus spécifique posée dans cette thèse, à savoir celle de l'effet d'une structure rythmique préalable sur la production de la parole. C'est pourquoi deux

cadres seront privilégiés ici. Le premier est PRISM (*Processing Rhythm in Speech and Music* ; Fiveash et al., 2021), qui propose une synthèse et des réflexions particulièrement utiles concernant les mécanismes potentiellement partagés entre traitement du rythme musical et traitement du rythme de la parole, offrant ainsi un cadre pour interpréter certains effets de l'amorçage rythmique. Le second est constitué par les modèles DIVA et GODIVA de la production de la parole (Bohland et al., 2010 ; Guenther, 2016 ; Tourville & Guenther, 2011), qui ne portent pas sur le rythme, mais offrent une architecture particulièrement pertinente pour penser la planification, le séquençage et le contrôle temporel de la parole. PRISM et DIVA/GODIVA éclairent ainsi des niveaux d'analyse différents : PRISM permet d'interroger les mécanismes communs au rythme musical et au rythme de la parole, notamment sur les plans perceptif, attentionnel et sensorimoteur, tandis que DIVA/GODIVA permettent de situer plus précisément ces effets dans une architecture de production de la parole.

3.1. Un cadre intégratif musique–parole pour penser l'amorçage rythmique

Le cadre PRISM ne repose pas sur l'idée que musique et parole seraient traitées par un même système. Il suggère plutôt que certains mécanismes impliqués dans le traitement du rythme peuvent être partagés entre les deux domaines, sans en effacer les spécificités propres (Fiveash et al., 2021). En cela, PRISM reprend et articule plusieurs lignes de travail antérieures portant sur les relations entre musique et langage, la prédiction temporelle, l'entraînement neuronal et le couplage sensorimoteur. Son intérêt, pour la présente thèse, est moins de fournir un modèle complet de la production de la parole que d'offrir un cadre pour penser comment une structure rythmique perçue en amont peut modifier les conditions temporelles du traitement puis, potentiellement, de la production verbale.

Dans cette perspective, PRISM met en avant trois grands mécanismes susceptibles d'être mobilisés, selon les tâches et les contextes, dans le traitement du rythme musical comme dans celui de la parole : (i) un traitement auditif précis permettant l'extraction fine des indices temporels et spectro-temporels ; (ii) la génération de prédictions temporelles via la synchronisation des oscillations neuronales à un rythme externe ; et (iii) un couplage

sensorimoteur reliant le traitement temporel à la préparation de l'action (Fiveash et al., 2021). Ces trois dimensions sont particulièrement utiles pour penser l'amorçage rythmique, dans la mesure où elles permettent d'articuler les propriétés du stimulus, les mécanismes temporels susceptibles d'être engagés, et la possibilité d'un effet ultérieur sur une tâche verbale (Figure 1).

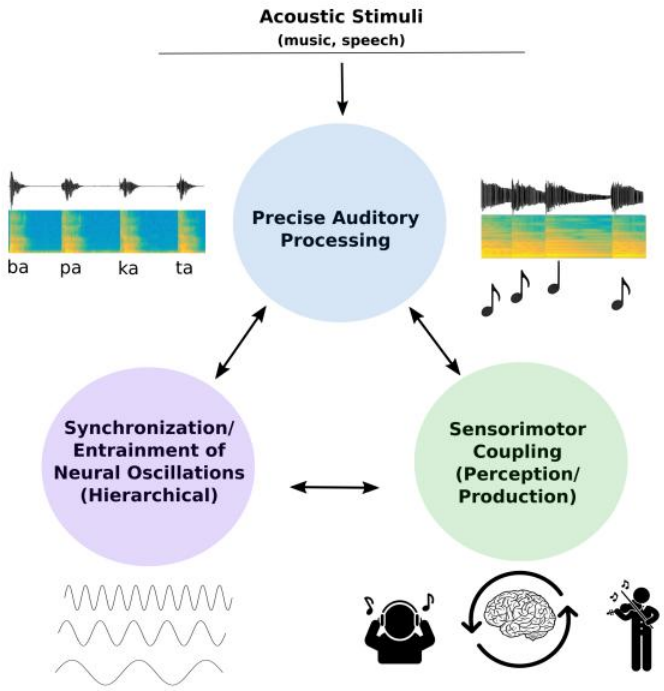

Figure 1 : Modèle PRISM. Figure adaptée de Fiveash et al., 2021.

La première composante concerne le traitement auditif précis (*fine-grained auditory processing*). Dans cette perspective, l'effet d'une amorce rythmique dépend d'abord de la capacité du système à encoder finement des indices acoustiques parfois brefs ou de faible amplitude, tels que les variations de durée, les attaques, les transitions spectrales ou les modulations lentes de l'enveloppe sonore. Comme le soulignent Kraus et Chandrasekaran (2010) et Patel (2011), cette précision perceptive est essentielle à la fois pour la discrimination phonémique et pour l'extraction d'indices temporels. Dans la parole, elle contribue à l'identification de contrastes segmentaux proches, à la segmentation syllabique, au calcul du débit et à certains processus de groupement prosodique ; en musique, elle participe au repérage des accents, du tempo et de la structure métrique. Peelle et Davis (2012) rappellent

d'ailleurs que les modulations lentes de l'enveloppe de parole véhiculent des informations importantes sur l'articulation, le débit et le rythme du signal. Cette première composante n'est pas spécifique à l'amorçage rythmique, mais elle en constitue une condition de possibilité : un stimulus temporellement structuré ne peut produire d'effet que dans la mesure où ses indices pertinents sont effectivement encodés.

La deuxième composante de PRISM concerne la prédiction temporelle et la synchronisation neuronale. Le rythme n'est pas seulement perçu : il est utilisé pour générer des attentes sur le moment probable de survenue des événements futurs. Cette idée a été formalisée dans la DAT, selon laquelle des dynamiques attentionnelles internes peuvent s'aligner sur la structure temporelle d'un signal externe, rendant certaines fenêtres temporelles plus saillantes et facilitant le traitement des événements attendus (Large & Jones, 1999). Dans le champ du traitement de la parole, ce principe a été développé par exemple par Peelle et Davis (2012), qui proposent que des oscillations neuronales de basse fréquence puissent suivre les modulations temporelles du signal de parole et soutenir ainsi sa compréhension. Dans cette perspective, l'amorce rythmique peut être conçue comme une mise en place préalable de conditions temporelles favorables : en exposant le système à une organisation temporelle régulière ou saillante, elle oriente en amont les mécanismes de prédiction susceptibles d'être mobilisés lors du traitement de la cible. Des travaux plus récents vont dans le même sens en suggérant que l'effet d'un rythme externe peut persister brièvement au-delà de la stimulation elle-même. Kösem et al. (2018) ont montré, en magnétoencéphalographie, que l'entraînement neuronal au rythme de parole précédent pouvait se maintenir pendant plusieurs cycles après un changement de débit, et que cette persistance influençait ensuite la perception d'un mot cible ambigu. Plus récemment, Zoefel et al. (2024) ont décrit, également en MEG, des *entrainment echoes* au niveau du cervelet, correspondant à des réponses rythmiques persistant brièvement après l'arrêt d'un stimulus de parole rythmiquement structuré. Les auteurs suggèrent que ces réponses pourraient refléter une mise à jour des prédictions temporelles après la fin du stimulus et leur relais vers des régions corticales impliquées dans le traitement de la parole. Ces résultats sont particulièrement importants pour l'amorçage rythmique, car ils rendent plausible l'idée qu'une structure temporelle entendue avant la cible puisse continuer à orienter, pendant un court laps de temps, les dynamiques attentionnelles et prédictives engagées par la tâche suivante. Dans cette optique, l'intérêt du rythme musical tient en grande partie à sa

prévisibilité temporelle souvent plus marquée que celle de la parole ordinaire. Alors que la parole présente une variabilité importante, liée au lexique, à la syntaxe, à la prosodie et aux objectifs communicatifs, la musique repose fréquemment sur une pulsation plus stable et une organisation métrique plus saillante. Cette différence ne signifie pas que la parole serait dépourvue de structure rythmique, mais que le rythme musical constitue souvent un contexte particulièrement favorable à l'installation d'attentes temporelles fortes.

Le troisième mécanisme mis en avant par PRISM est le couplage sensorimoteur. Fiveash et al. (2021) soulignent que le traitement du rythme, en particulier lorsqu'il repose sur une régularité suffisante pour permettre l'anticipation, s'accompagne souvent d'interactions étroites entre perception et motricité. Dans le domaine musical, Patel et Iversen (2014), avec l'hypothèse ASAP, proposent que la perception du *beat* engage des régions de planification motrice, même en l'absence de mouvement manifeste. L'idée n'est pas que la perception du rythme soit réductible à une motricité cachée, mais que les systèmes moteurs puissent contribuer à la génération de prédictions temporelles utiles pour l'écoute. Dans le cadre de PRISM, cette composante ne porte pas directement sur la production de la parole ; elle suggère plutôt qu'un rythme externe peut être d'autant plus efficace qu'il recrute, au-delà de la seule perception, des ressources de préparation de l'action. Cette lecture est cohérente avec plusieurs résultats expérimentaux. Cason et al. (2015a) ont montré que des amorces métriques pouvaient améliorer la perception de la parole et que cet effet était amplifié par un entraînement audio-moteur préalable. De même, Falk et Dalla Bella (2016) ainsi que Falk et al. (2017) suggèrent que l'alignement entre rythmes de parole, tappings moteurs et structure de la cible peut améliorer certaines performances de traitement verbal. Ces travaux ne démontrent pas que tout effet d'amorçage soit nécessairement d'origine sensorimotrice, mais ils indiquent qu'un rythme externe peut gagner en efficacité lorsqu'il s'inscrit dans une dynamique perception-action plus étroite. Dans cette perspective, la synchronisation sensorimotrice, mesurée notamment à partir de tâches de tapping, peut être envisagée comme un indicateur intégratif reliant extraction du rythme, prédiction temporelle et couplage perception-action. Dans le cadre de PRISM, des performances de tapping plus stables ou plus précises ne doivent donc pas être interprétées uniquement comme des performances motrices, mais comme des indices possibles de la capacité à extraire une structure rythmique, à générer des attentes temporelles stables et à les mobiliser dans la préparation de l'action (Fiveash et al., 2021).

PRISM fournit ainsi un cadre particulièrement utile pour penser l'amorçage rythmique comme une exposition préalable susceptible de préparer temporellement le système. Une amorce régulière ou musicalement structurée peut, en théorie, agir à plusieurs niveaux : en facilitant l'encodage auditif des indices temporels, en renforçant la prédiction via l'entraînement neuronal et, dans certaines conditions, en mobilisant plus efficacement le couplage perception–action. Toutefois, PRISM ne précise pas directement à quel niveau de la production de la parole un tel effet pourrait s'inscrire. Cette limite justifie de se tourner vers des modèles plus directement centrés sur la production, afin de mieux caractériser les étapes susceptibles d'être influencées par ce type d'amorçage.

3.2. Modèles de production de la parole

Plusieurs modèles ont été proposés dans la littérature pour rendre compte de la production de la parole. Les modèles psycholinguistiques ont surtout permis de décrire les étapes de l'encodage verbal, depuis la conceptualisation jusqu'à la programmation phonologique et phonétique (Levelt et al., 1999), ou encore les dynamiques d'activation interactive entre unités linguistiques (Dell, 1986). À côté de ces approches centrées sur l'encodage linguistique, d'autres cadres se sont davantage intéressés au contrôle moteur et sensorimoteur de la parole. Ces modèles de contrôle moteur se distinguent notamment par la place qu'ils accordent aux commandes motrices anticipatrices, aux boucles de rétroaction sensorielle et à l'estimation de l'état du système de production (Parrell et al., 2019). C'est dans cette perspective que s'inscrivent, par exemple, le modèle de *State Feedback Control* de Houde et Nagarajan (2011), centré sur le rôle des prédictions sensorielles dans l'ajustement de la production, ou encore le modèle HSFC (*Hierarchical State Feedback Control*) de Hickok (2012), qui propose une organisation plus générale des relations entre perception et production verbales.

Dans le cadre de cette thèse, ces différents modèles ne présentent toutefois pas tous la même utilité. Les modèles psycholinguistiques rendent compte avec précision de la préparation linguistique de la parole, mais ils spécifient moins directement les mécanismes moteurs et sensorimoteurs susceptibles d'être modulés par une préparation temporelle externe. À cet égard, les modèles DIVA (*Directions Into Velocities of Articulators*) et GODIVA

(*Gradient Order Directions Into Velocities of Articulators*) apparaissent particulièrement pertinents, car ils constituent des modèles neurocomputationnels de la production de la parole qui relient représentations de parole, commandes motrices, boucles de feedback et organisation séquentielle à des réseaux cérébraux identifiables (Bohland et al., 2010 ; Tourville & Guenther, 2011 ; Guenther, 2016 ; Miller & Guenther, 2021). Ils ne modélisent pas directement l'effet d'une stimulation rythmique externe, ni l'influence d'un interlocuteur sur la production, mais ils permettent de formuler des hypothèses sur les composantes de la production qui pourraient être sensibles à une modulation temporelle préalable. C'est en cela qu'ils sont particulièrement adaptés à la présente thèse.

3.2.1. DIVA

Le modèle DIVA (*Directions Into Velocities of Articulators*) a été élaboré pour rendre compte des mécanismes cérébraux impliqués dans la production de la parole à partir d'une articulation entre commandes feedforward et contrôles feedback auditif et somatosensoriel (Tourville & Guenther, 2011, voir Figure 2). Son idée centrale est que la parole ne repose pas seulement sur l'exécution de gestes articulatoires, mais sur l'activation de programmes moteurs appris, guidés par des attentes sensorielles et ajustés, si nécessaire, à partir des écarts entre les conséquences attendues et les conséquences effectivement perçues. Dans cette perspective, parler revient à mobiliser un système prédictif, dans lequel l'action et la perception sont étroitement couplées.

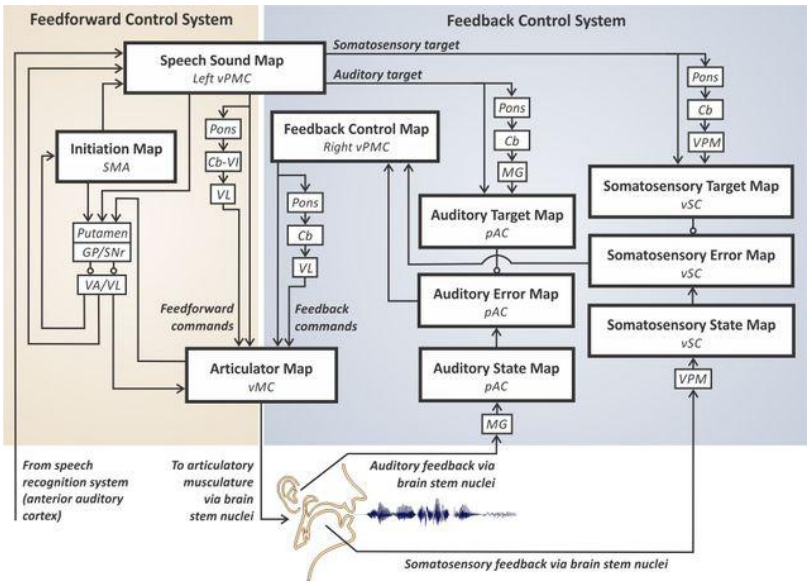

Figure 2 : Modèle DIVA. Figure adaptée de Guenther, 2016.

Au centre de l'architecture se trouve la *speech sound map*, classiquement associée au cortex prémoteur ventral gauche. Cette carte contient des représentations de sons de parole appris, pouvant correspondre à des unités relativement stabilisées de la production, souvent décrites comme des *chunks* sensorimoteurs (Guenther et al., 2006 ; Tourville & Guenther, 2011). L'activation d'un de ces nœuds entraîne la lecture de commandes motrices vers le cortex moteur primaire, directement et via une boucle ponto-cérébello-thalamique. Le système feedforward constitue ainsi la composante dominante de la parole fluente : il permet de produire rapidement et efficacement les gestes nécessaires sans dépendre à chaque instant du retour sensoriel.

Cette prépondérance du feedforward n'implique cependant pas l'absence de contrôle sensoriel. DIVA accorde une place essentielle au feedback auditif et au feedback somatosensoriel. Dans le premier cas, le modèle postule l'existence d'une *auditory target map*, qui code les conséquences auditives attendues de la production en cours, et d'une *auditory state map*, qui représente l'état auditif effectivement reçu. Si un écart apparaît entre les deux, une *auditory error map* s'active, puis transmet cette erreur à une *feedback control map*, qui la convertit en commande motrice correctrice. Une logique parallèle vaut pour le système somatosensoriel, qui compare les états tactiles et proprioceptifs attendus et observés du tractus vocal. La parole apparaît donc, dans DIVA, comme une activité contrôlée par des cibles sensorielles internalisées, et non comme un simple enchaînement de gestes

articulatoires. Le feedback n'y est pas seulement conçu comme un mécanisme de réparation de l'erreur ; il joue également un rôle fondamental dans l'apprentissage et le recalibrage progressif des commandes feedforward. C'est en ce sens que DIVA constitue un modèle de la parole à la fois comme performance motrice et comme système d'apprentissage sensorimoteur.

Dans ses développements ultérieurs, le cadre DIVA a également intégré plus explicitement la question du déclenchement des programmes moteurs, en attribuant un rôle particulier à une *initiation map* située dans la SMA, en interaction avec les ganglions de la base et le thalamus (Guenther, 2016). Cette précision est importante, car elle permet déjà de distinguer, au sein du système de parole, la représentation d'un programme moteur, sa libération effective et son contrôle sensoriel. DIVA doit ainsi être compris avant tout comme un modèle du contrôle sensorimoteur des unités de parole apprises, particulièrement puissant pour décrire la manière dont un son ou une séquence familière est produit, surveillé et ajusté.

3.2.2. GODIVA

Si DIVA rend compte avec précision de l'exécution et du contrôle d'unités de parole apprises, il ne suffit pas, à lui seul, à expliquer comment plusieurs unités sont préparées, maintenues et ordonnées dans des énoncés plus longs. Or, la parole ordinaire implique précisément qu'un locuteur puisse activer simultanément plusieurs éléments à venir, tout en les produisant dans le bon ordre. C'est pour traiter cette dimension sérielle qu'a été développé GODIVA, conçu comme une extension de DIVA au séquençage de la parole (Bohland et al., 2010, voir Figure 3).

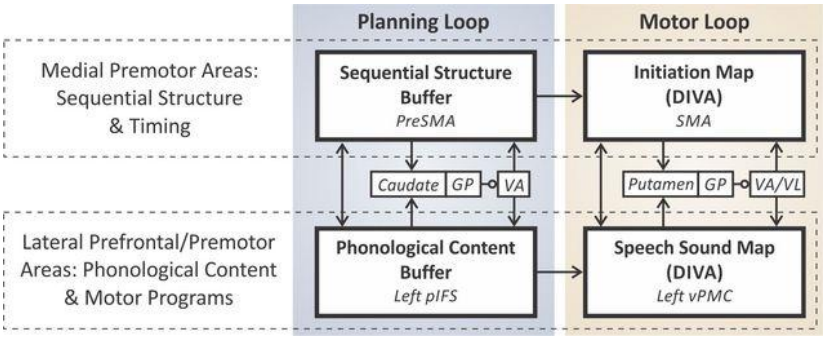

Figure 3 : Modèle GODIVA. Figure adaptée de Guenther, 2016.

L'apport principal de GODIVA est de distinguer explicitement le contenu phonologique d'un énoncé et sa structure séquentielle. Le contenu phonologique est maintenu dans un phonological content buffer, localisé dans le *posterior inferior frontal sulcus* gauche, tandis que l'ordre de ces éléments et la structure plus générale de la séquence sont représentés dans un sequential structure buffer, localisé dans la pré-SMA (Bohland et al., 2010 ; Guenther, 2016). Cette distinction est décisive : elle revient à poser que préparer un énoncé ne consiste pas seulement à activer une série d'unités phonologiques, mais aussi à spécifier leur ordre de déploiement.

Pour rendre compte de cet ordonnancement, GODIVA mobilise un mécanisme de *competitive queuing*. Plusieurs unités sont activées en parallèle, mais avec un gradient d'activation qui reflète leur rang dans la séquence. L'élément prévu en premier est le plus activé ; les suivants le sont à un degré moindre. Au moment où le premier élément est sélectionné pour l'exécution, son activation décroît, ce qui permet au suivant de devenir dominant. L'ordre sériel ne dépend donc pas d'un simple chaînage linéaire, mais d'une dynamique de compétition organisée entre plusieurs représentations temporairement coactives. Cette proposition permet de penser plus finement la fluidité de la parole connectée, mais aussi les erreurs d'ordre ou les difficultés de séquençage observées dans certains troubles de la parole.

GODIVA distingue en outre une boucle de planification et une boucle motrice. La première implique les régions préfrontales et médiales nécessaires au maintien du contenu phonologique et de la structure séquentielle ; la seconde rejoint les composantes plus directement motrices déjà décrites dans DIVA. Les ganglions de la base y occupent une position importante, dans la mesure où ils participent au *gating* des représentations, c'est-à-dire au passage du plan d'énoncé vers le programme moteur effectivement mis en route. L'intérêt théorique de GODIVA est ainsi de fournir une interface explicite entre mémoire de travail phonologique, organisation séquentielle et exécution motrice.

Là où DIVA éclaire surtout la parole comme action sensorimotrice prédictive appliquée à des unités déjà disponibles, GODIVA permet donc de penser la parole connectée, la préparation de séquences et le passage ordonné d'une unité à la suivante. Cette complémentarité est essentielle : DIVA fournit le socle du contrôle sensorimoteur ; GODIVA

en étend la portée à la dimension sérielle, indispensable à la production des énoncés de plusieurs syllabes, plusieurs mots.

3.2.3. DIVA/GODIVA et la maladie de Parkinson

C'est dans l'application de ces modèles à la MP que l'article de Manes et al. (2024) devient particulièrement utile, illustrant comment ce cadre peut être mobilisé pour relire les troubles de la parole dans la MP à partir du rôle des ganglions de la base dans l'initiation, le gain moteur, l'apprentissage séquentiel et le séquençage de la parole continue. Les auteurs défendent une position nuancée : les ganglions de la base ne seraient pas les seuls garants de la parole adulte, ni même indispensables à tous ses aspects, mais ils contribueraient à plusieurs opérations importantes dont l'altération permet d'éclairer une partie du profil dysarthrique observé dans la maladie.

Le premier point concerne l'initiation et la mise à l'échelle des mouvements de parole. Dans la lecture proposée par Manes et al. (2024), la déplétion dopaminergique réduit le soutien excitateur apporté par la boucle cortico-basal-ganglio-thalamo-corticale au nœud correspondant au programme moteur courant dans l'*initiation map* de la SMA. Cette réduction du soutien aurait au moins deux conséquences. D'une part, le programme moteur en cours serait moins fortement activé, ce qui contribuerait à la diminution du gain des mouvements de parole et, par là, à des phénomènes tels que l'hypophonie ou la réduction de l'ampleur articulatoire. D'autre part, le programme courant deviendrait plus vulnérable à une terminaison prématurée, parce qu'il résisterait moins bien à la compétition exercée par les programmes suivants dans la séquence. Dans cette perspective, la parole parkinsonienne ne résulte pas seulement d'une faiblesse motrice diffuse, mais aussi d'une instabilité dans le maintien du programme en cours. Cette hypothèse présente l'intérêt de rendre compte d'observations cliniques qui pourraient paraître contradictoires. Elle permet notamment de comprendre pourquoi certains profils de parole parkinsonienne associent une réduction de l'intensité vocale à une accélération du débit. Dans ce cadre, l'accélération n'est pas l'inverse d'un déficit moteur ; elle peut au contraire refléter une difficulté croissante à maintenir activé le programme courant, conduisant le système à passer trop rapidement au suivant. Ainsi, la réduction du gain moteur et l'accélération de la parole peuvent être interprétées comme deux

effets possibles d'un même déséquilibre dans le soutien du programme moteur par les ganglions de la base.

Le deuxième apport du cadre DIVA/GODIVA concerne l'apprentissage de nouvelles séquences de parole. Dans la perspective défendue par Manes et al. (2024), les ganglions de la base joueraient un rôle important dans l'apprentissage initial de séquences motrices, y compris verbales, tandis que la production de séquences très bien apprises dépendrait davantage de circuits corticaux. Une telle hypothèse permet d'expliquer pourquoi la MP n'entraîne pas typiquement d'erreurs massives d'ordre dans les séquences surapprises, tout en restant compatible avec des difficultés dans l'acquisition de nouvelles séquences phonologiques ou verbales. Le cadre invite ainsi à distinguer clairement la relative préservation des séquences familières et la fragilité potentielle de l'apprentissage de séquences nouvelles. Le troisième point concerne le buffering temporaire et le séquençage des énoncés plus longs. Sur ce plan, Manes et al. (2024) défendent une position prudente. Les ganglions de la base pourraient contribuer au soutien du programme courant et au passage d'une unité à la suivante dans la parole continue, notamment lorsque les énoncés sont longs ou peu routinisés. En revanche, la littérature ne plaide pas pour un rôle critique des ganglions de la base dans la représentation de l'ordre sériel lui-même. L'absence d'erreurs d'ordre caractéristiques dans la parole parkinsonienne et la relative préservation de la parole après pallidotomie bilatérale, chirurgie lésionnelle du globus pallidus utilisée pour traiter la MP, suggèrent plutôt que leur rôle serait plutôt facilitateur que strictement obligatoire. Dans cette perspective, les ganglions de la base participeraient moins à représenter l'ordre abstrait des séquences qu'à soutenir leur déploiement dynamique, leur intensité et leur maintien temporel.

Manes et al. (2024) discutent également les données relatives aux feedbacks auditif et somatosensoriel. Leur lecture demeure globalement compatible avec l'architecture de DIVA, selon laquelle les ganglions de la base n'interviennent pas directement dans les circuits de feedback auditif ou somatosensoriel. Cela permet de comprendre pourquoi plusieurs études rapportent des réponses réflexes auditives préservées, voire exagérées, chez les personnes atteintes de la MP, alors même que l'adaptation sensorimotrice à plus long terme peut être réduite. Dans ce cadre, les anomalies observées renverraient moins à un déficit primaire du feedback lui-même qu'à un déséquilibre entre systèmes feedforward et feedback, ou à une

difficulté de mise à jour durable des commandes feedforward à partir de l'erreur sensorielle. Ce point constitue d'ailleurs l'une des limites reconnues du cadre actuel : la contribution exacte de la dopamine et des boucles cortico-ganglio-basal à l'adaptation auditivo-motrice n'y est pas encore entièrement spécifiée.

Enfin, l'application du cadre posé par DIVA/GODIVA à la MP permet aussi d'éclairer les effets thérapeutiques. L'hétérogénéité des effets de la dopamine sur la parole, la grande variabilité observée avec la stimulation cérébrale profonde, ou encore la relative préservation de certaines dimensions de la parole après pallidotomie, deviennent plus lisibles si l'on admet que plusieurs composantes distinctes de la production peuvent être différemment affectées : le gain moteur, l'initiation, le maintien du programme courant, l'apprentissage séquentiel ou le contrôle sensorimoteur. En ce sens, l'intérêt principal de DIVA et de GODIVA pour l'étude de la MP n'est pas d'offrir une explication totale des troubles, mais de permettre une décomposition mécaniste de leurs différentes facettes. C'est précisément ce qui en fait un cadre particulièrement fécond pour un manuscrit de thèse consacré à la parole pathologique.

Dans cette thèse, PRISM et DIVA/GODIVA ne sont pas mobilisés parce qu'ils décriraient un même mécanisme, mais parce qu'ils permettent de situer l'effet possible du rythme à des niveaux d'analyse différents. PRISM aide à comprendre comment une structure rythmique perçue peut préparer temporellement le système, en agissant sur l'encodage auditif, la prédiction temporelle et le couplage perception–action. DIVA et GODIVA permettent, quant à eux, d'interroger les étapes de la production où cette préparation pourrait se traduire : initiation, activation du programme moteur, ordonnancement des unités, maintien de la séquence en cours, contrôle feedforward et ajustements sensorimoteurs. L'amorçage rythmique est ainsi envisagé non comme une action globale sur la parole, mais comme une modulation préalable dont les effets peuvent se manifester sur certaines composantes spécifiques de la production.

Chapitre 4
Problématique et objectifs

Ainsi, les troubles de la parole dans la MP ne se réduisent ni à une simple altération de la voix ni à une atteinte périphérique de l'articulation. Ils concernent plus largement l'organisation temporelle et prosodique de la production, avec des perturbations touchant notamment l'initiation, le débit, l'accentuation, la modulation mélodique et la structuration des pauses, autant de dimensions qui participent directement à l'intelligibilité et à l'efficacité communicative (Darley et al., 1969a, 1969b, 1969c, 1975 ; Duffy, 2020). Dans ce contexte, l'hypothèse selon laquelle un rythme auditif externe pourrait influencer la parole présente un intérêt particulier, à la fois pour mieux comprendre les mécanismes du contrôle temporel de la production et pour envisager de nouvelles formes de soutien en clinique.

Les éléments présentés précédemment suggèrent que l'amorçage rythmique constitue une piste pertinente pour soutenir la parole dans la MP, à la fois sur le plan théorique, à travers les cadres PRISM et DIVA/GODIVA, et sur le plan clinique, à travers les interventions fondées sur le rythme. Toutefois, plusieurs incertitudes majeures demeurent quant aux dimensions de parole concernées, aux mécanismes impliqués et aux conditions dans lesquelles un tel effet peut être observé.

La première concerne les dimensions de parole réellement sensibles à l'amorçage rythmique. La littérature sur la prosodie dans la MP reste en effet marquée par une forte hétérogénéité, y compris pour des paramètres considérés comme centraux, tels que l'étendue des variations de F0, certaines mesures temporelles ou divers indices de rythme. Cette hétérogénéité complique l'identification de cibles robustes et limite l'interprétation des effets d'amorçage : en l'absence d'un cadre suffisamment clair sur les altérations les plus stables, il devient difficile de déterminer ce que signifie un effet positif, nul ou contrasté.

La deuxième incertitude porte sur la nature même de l'effet d'amorçage. Les travaux disponibles suggèrent qu'un indice temporel externe peut parfois faciliter certains aspects de

la production, mais il reste mal établi si cet effet dépend avant tout de la régularité du stimulus, de sa structure interne, de sa modalité, ou encore des propriétés de la tâche de parole elle-même. En outre, les études consacrées à la MP reposent encore largement sur des amorces verbales, tandis que les effets d'amorces musicales et leurs conséquences sur des dimensions explicitement prosodiques demeurent relativement peu documentés (Späth et al., 2016 ; Aichert et al., 2021). Il reste donc à préciser si l'amorçage agit principalement sur des dimensions globales de préparation temporelle, telles que l'initiation ou le débit, ou s'il peut également affecter des niveaux plus fins d'organisation prosodique.

La troisième incertitude concerne les mécanismes susceptibles de sous-tendre l'effet observé. Les notions d'entraînement neuronal, de prédiction temporelle et de couplage sensorimoteur sont fréquemment mobilisées pour interpréter les bénéfices potentiels d'un rythme externe, mais elles restent encore trop rarement mises à l'épreuve dans des dispositifs articulant simultanément mesures neurophysiologiques, indices de synchronisation sensorimotrice et mesures acoustiques de la parole. Autrement dit, plusieurs mécanismes plausibles ont été proposés, mais leur contribution effective à la modulation de la parole demeure encore insuffisamment contrainte empiriquement. Cette limite freine non seulement la compréhension théorique du phénomène, mais aussi l'identification de marqueurs susceptibles de prédire quels individus seront les plus sensibles à l'amorçage.

Enfin, une quatrième incertitude concerne la variabilité de l'effet. Si un soutien rythmique externe n'agit pas de manière uniforme selon les individus ou les populations, la question ne peut pas être réduite à savoir s'il « fonctionne » ou non en moyenne. Elle doit également porter sur les conditions dans lesquelles il est susceptible d'être efficace, sur les profils qui en bénéficient le plus, et sur les réseaux neurofonctionnels impliqués. Dans cette perspective, la comparaison entre personnes atteintes de MP, personnes atteintes d'ataxie cérébelleuse et participants contrôles constitue un levier particulièrement important. Elle permet d'interroger, de manière indirecte mais informative, la contribution relative des circuits associés aux ganglions de la base et de ceux impliquant le cervelet dans le traitement temporel et le contrôle de la parole.

Dans ce contexte, les objectifs de cette thèse sont les suivants :

Le premier objectif est d'identifier les dimensions prosodiques et temporelles de la parole les plus robustement altérées dans la MP. Avant d'examiner les effets possibles d'un amorçage rythmique, il apparaît en effet nécessaire de préciser quelles dimensions de la parole constituent des cibles pertinentes, à la fois sur le plan théorique et méthodologique. Cet objectif vise ainsi à mieux cerner, à partir de la littérature existante, les paramètres acoustiques les plus sensibles aux troubles prosodiques dans la MP, en tenant compte de l'hétérogénéité importante des résultats disponibles.

Le deuxième objectif consiste à déterminer si l'exposition à une amorce rythmique auditive modifie effectivement la production de parole ultérieure, et à identifier les dimensions les plus sensibles à cette modulation. Une attention particulière est portée au temps d'initiation, au débit, ainsi qu'à différents indices de structuration temporelle et prosodique. Il s'agit ainsi d'examiner si l'effet de l'amorçage concerne principalement la préparation temporelle globale de la parole, ou s'il peut également s'étendre à des niveaux plus fins d'organisation de la parole.

Le troisième objectif repose sur la comparaison entre personnes atteintes de MP, personnes atteintes d'ataxie cérébelleuse et participants contrôles. Cette comparaison doit permettre de mieux contraindre l'interprétation mécanistique des effets observés, en examinant si certaines composantes de l'amorçage rythmique sont plus particulièrement sensibles à l'intégrité des circuits impliquant les ganglions de la base ou le cervelet. Dans la MP, l'altération des circuits intégrant les ganglions de la base et des mécanismes de génération interne du rythme pourrait rendre les individus plus dépendants d'un soutien temporel externe. À l'inverse, l'atteinte cérébelleuse pourrait limiter certains ajustements temporels fins nécessaires à l'exploitation efficace de ce soutien. La comparaison entre ces profils neurofonctionnels constitue ainsi un moyen indirect mais informatif d'interroger les mécanismes compensatoires susceptibles d'être mobilisés par l'amorçage rythmique.

Le quatrième objectif est d'examiner si les effets observés en production seraient en lien avec des indices compatibles de prédictions temporelles et un couplage perception–action. Cet objectif vise ainsi à dépasser une lecture strictement descriptive du paradigme, en examinant dans quelle mesure les réponses à l'amorçage rythmique peuvent être associées à des mesures de synchronisation sensorimotrice et de compétences musicales.

Enfin, le cinquième objectif est d'interroger l'intérêt clinique des approches basées sur le rythme et leur place possible dans la pratique orthophonique actuelle. Il s'agit, d'une part, de replacer les résultats expérimentaux de la thèse dans une perspective de pratique fondée sur les preuves, et, d'autre part, de mieux comprendre comment les orthophonistes perçoivent, utilisent ou pourraient intégrer les activités musicales et rythmiques dans leurs prises en soin.

Afin de répondre aux questions de recherche et aux objectifs exposés précédemment, la thèse s'appuie sur un ensemble de travaux présentés dans la seconde partie du manuscrit sous la forme d'articles scientifiques publiés, soumis ou en préparation.

Le premier article correspond à une méta-analyse consacrée aux altérations acoustiques de la prosodie dans la MP. Il ouvre le programme de recherche en précisant quelles dimensions prosodiques apparaissent les plus robustement affectées dans cette population, avant d'examiner expérimentalement leur possible modulation par l'amorçage rythmique. En synthétisant une littérature marquée par une forte hétérogénéité, cette méta-analyse permet d'identifier les paramètres de parole les plus pertinents pour les études expérimentales de la thèse. Elle répond donc directement au premier objectif, qui consiste à clarifier les dimensions de parole à privilégier pour étudier les effets potentiels d'un soutien rythmique externe.

Le deuxième article, constitue le premier volet expérimental. Il examine les effets de l'amorçage rythmique sur l'initiation de la parole et sur des indices plus globaux de structuration temporelle et rythmique, en comparant des personnes atteintes de la MP, des personnes atteintes d'ASCS et des participants contrôles. Cette étude répond directement aux

objectifs d'évaluer l'impact de l'amorçage rythmique sur la production de la parole et d'examiner la contribution relative de profils neurofonctionnels distincts.

Le troisième chapitre est le prolongement de l'article précédent. Il vise à déterminer si l'amorçage rythmique agit uniquement sur des paramètres globaux de temporalité ou s'il est également susceptible d'affecter des dimensions tels que la F0 et les pauses. Ce troisième volet contribue ainsi à préciser quelles dimensions de la parole sont effectivement sensibles à l'amorçage, et à quel niveau d'organisation cette modulation peut se manifester.

Le quatrième chapitre apporte un éclairage complémentaire sur les mécanismes susceptibles de sous-tendre les effets observés. Réalisée auprès d'adultes tout-venant, cette étude combine mesures comportementales, et indices de synchronisation sensorimotrice afin d'examiner dans quelle mesure l'effet d'un amorçage rythmique peut être relié au couplage perception–action.

Le cinquième chapitre, et troisième article, ouvre enfin la réflexion vers les pratiques orthophoniques. En documentant les usages, les représentations et les modalités d'intégration du rythme et de la musique dans les pratiques des orthophonistes francophones, il apporte un prolongement plus directement translationnel aux travaux expérimentaux de la thèse et permet de replacer les résultats obtenus dans une perspective clinique plus large. Il contribue ainsi à l'objectif consistant à interroger les conditions de transférabilité du paradigme dans une logique orthophonique fondée sur les preuves.

Deuxième partie

Contributions et études expérimentales

Chapitre 1
Mesures acoustiques de la prosodie dans la maladie de Parkinson : revue systématique et méta-analyse bayésienne

Si la dysprosodie est décrite depuis longtemps comme une composante importante de la dysarthrie hypokinétique, les données disponibles demeurent hétérogènes. Selon les études, les altérations rapportées ne concernent pas toujours les mêmes paramètres, et leur ampleur varie selon les tâches, les langues ou les caractéristiques des participants. Dans ce contexte, une synthèse systématique de la littérature semblait nécessaire afin de distinguer les marqueurs prosodiques les plus solidement associés à la MP de ceux pour lesquels les résultats restent plus incertains. C'est précisément l'objectif d'une des études réalisées dans cette thèse. Il s'agit d'une revue systématique de la littérature et une méta-analyse, publiée dans la revue *Journal of Speech, Language, and Hearing Research en 2024*, qui est présentée sous-forme d'un article dans ce chapitre.

En résumé, l'article est intitulé *A systematic review and Bayesian meta-analysis of acoustic measures of prosody in Parkinson's disease* et porte sur 36 articles sélectionnés à partir de 648 références initiales. Il s'intéresse en particulier à la variation de la F0, à la variation d'intensité, au débit de parole, au débit articulatoire, ainsi qu'aux caractéristiques des pauses. Au-delà de la synthèse quantitative qu'il fournit, ce travail présente un intérêt central en ce qu'il permet d'identifier les dimensions prosodiques les plus robustement affectées dans des tâches de parole quand on s'intéresse à la dysprosodie dans la MP.

Les résultats mettent plus particulièrement en évidence deux paramètres apparaissant comme les plus altérés dans la MP : la réduction de la variabilité de F0 et l'augmentation de la durée des pauses. À l'inverse, d'autres dimensions souvent évoquées dans la littérature clinique apparaissent moins robustes.

Le résumé en français et l'article dans sa version publiée sont présentés ci-après.

Fumel, J., Bahuaud, D., Weed, E., Fusaroli, R., & Basirat, A. (2024). A Systematic Review and Bayesian Meta-Analysis of Acoustic Measures of Prosody in Parkinson's Disease. *Journal of speech, language, and hearing research : JSLHR*, *67*(8), 2548–2564. https://doi.org/10.1044/2024_JSLHR-23-00588

Résumé

Objectif : la prosodie linguistique est altérée dans la maladie de Parkinson (MP), ce qui suggère un rôle des ganglions de la base dans la production de la prosodie. Cependant, il n'existe aucune synthèse systématique récente des données acoustiques disponibles concernant les troubles prosodiques dans la MP. Cette étude visait à identifier les caractéristiques acoustiques de la prosodie linguistique qui sont systématiquement affectées dans la MP.

Méthode : la revue systématique des articles rend compte des caractéristiques acoustiques de la production prosodique dans la MP. Les articles portaient sur la fréquence fondamentale (F0) et sa variabilité, l'intensité et sa variabilité, le débit de parole et le rythme d'articulation, ainsi que la durée et le ratio des pauses. Sur un total de 648 études identifiées, 36 répondaient aux critères d'inclusion et d'exclusion. Pour chaque mesure acoustique et chaque tâche, les données des personnes atteintes de la MP ont été comparées à celles d'un groupe témoin afin d'extraire les tailles d'effet. Les tailles d'effet regroupées ont été estimées à l'aide de modèles de régression hiérarchique bayésiens robustes.

Résultats : la MP était associée à une diminution de la variabilité de la F0 et à une augmentation de la durée des pauses. Les preuves d'une réduction de la variabilité de l'intensité et du débit de parole chez les personnes atteintes de la MP étaient limitées. Aucune preuve n'a été trouvée suggérant que la MP affecte le débit d'articulation ou le ratio de pause.

Conclusions : les principaux paramètres acoustiques de la prosodie affectés par la MP sont la variabilité de la F0 et la durée des pauses. L'identification de ces paramètres acoustiques a d'importantes implications cliniques pour le choix des stratégies de prise en charge de la MP. L'association entre la variabilité de la F0 et la durée des pauses avec la MP suggère que les circuits neuronaux contrôlant ces paramètres sont au moins partiellement communs et pourraient inclure les ganglions de la base. Alors que la présente étude s'est concentrée sur la réalisation phonétique des indices prosodiques, de futures études devraient examiner si et comment la MP affecte la prosodie à des niveaux de traitement plus élevés.

Mots-clés : prosodie, maladie de Parkinson, acoustique, parole

A systematic review and Bayesian meta-analysis of acoustic measures of prosody in Parkinson's disease

Jules Fumel[1], Delphine Bahuaud[2], Ethan Weed[3,4], Riccardo Fusaroli[3,4], Anahita Basirat[1]

[1] Univ. Lille, CNRS, UMR 9193 SCALab, F-59000 Lille, France

[2] Department of Speech and Language Therapy, Faculty of Medicine, UFR3S, F-59000 Lille, France

[3] Linguistics, Cognitive Science and Semiotics, School of Communication and Culture, Aarhus University, Aarhus, Denmark

[4] Interacting Minds Center, School of Culture and Society, Aarhus University, Aarhus, Denmark

Abstract

Purpose: Linguistic prosody is affected in Parkinson's Disease (PD), which implicates the basal ganglia's role in the production of prosody. However, there is no recent systematic synthesis of the available acoustic evidence of prosodic impairment in PD. This study aimed to identify the acoustic features of linguistic prosody that are consistently affected in PD.

Method: The authors systematically reviewed papers that reported acoustic features of prosodic production in PD. Articles focused on fundamental frequency (F0) and its variability, intensity and its variability, speech and articulation rate, and pause duration and ratio. From a total of 648 records identified, 36 met criteria for inclusion and exclusion. For each acoustic measurement and task, data from people with PD (PwPD) were compared with controls to extract effect sizes. Pooled effect sizes were estimated using robust Bayesian hierarchical regression models.

Results: PD was associated with decreased F0 variability and increased pause duration. There was limited evidence of reduced intensity variability and speech rate in PwPD. No evidence was found to suggest that PD affects articulation rate or pause ratio.

Conclusions: The primary acoustic parameters of prosody affected by PD are F0 variability and pause duration. The identification of these acoustic parameters has important clinical implications for the selection of PD management strategies. The association of F0 variability and pause duration with PD suggests that the neural circuits controlling these parameters are at least partly shared and might include the basal ganglia. While the current study focused on the phonetic realization of prosodic cues, future studies should examine whether and how PD affects prosody at higher levels of processing.

Keywords: Prosody, Parkinson's disease, acoustic, speech

Introduction

Parkinson's disease (PD) leads to changes in the communicative abilities of most people with Parkinson's disease (PwPD). Prosody is one of the main dimensions of communication reported as impaired in PD (Miller, 2017). Prosody is an umbrella term which refers to the aspects of speech that go beyond segments (i.e., vowels and consonants), such as intonation and accentuation. Prosody conveys linguistic cues such as focus, sentence mode and phrasal boundaries, as well as paralinguistic information including speaker's attitudes and emotions (Van Lancker Sidtis & Yang, 2020) and is considered one of the crucial aspects of natural interactive language (Levinson, 2019). Although the presence of prosodic impairment in PD has been known for decades and while 'dysprosody' is a part of the perceptual description of hypokinetic dysarthria as defined by Darley et al. (1969), much remains to be understood about the abilities of PwPD to use prosody to convey meaning (Frota et al., 2021), the mechanisms which underlie prosody impairments (Arnold et al., 2014), and the possibilities to rehabilitate the prosodic function in PwPD (Atkinson-Clement et al., 2015). This study aimed to enhance our understanding of the impact of PD on the acoustic features of prosody through a systematic review and meta-analysis.

Prosody impairment in PD is traditionally examined using two approaches: perceptual assessment and acoustic analysis. Perceptual assessment involves evaluating how listeners perceive the prosody produced by speakers. In such studies, naive participants or trained evaluators are asked to rate the utterances produced by PwPD (e.g., lexical stress, Pell et al., 2006). Acoustic analysis investigates key acoustic features of prosody quantitatively, i.e., fundamental frequency (F0), amplitude and duration (e.g., the rise and fall of F0 in statements vs. questions, Basirat et al., 2018). While these approaches are complementary, acoustic analysis is an increasingly popular choice as it can be implemented in tools to automatically assess and monitor speech impairment in PD (Laganaro et al., 2021), including remote telemedical tools (Kouba et al., 2022). A growing body of research focuses on speech-based biomarkers of PD using acoustic analysis (e.g., Rusz, Hlavnička, et al., 2021), including machine-learning approaches (e.g., Arora et al., 2019; Jeancolas et al., 2021). Although promising, the diagnostic accuracy of these analyses in terms of sensitivity (true-positive rate) and specificity (true-negative rate) remains unclear (Sharma et al., 2023). To improve the accuracy of these analyses, it is essential to incorporate knowledge about various aspects of speech production

impairments into the data processing pipeline (Ramanarayanan et al., 2022). Meta-analyses play an important role in providing this crucial knowledge. In sum, a better understanding of how PD affects the acoustic features of speech is crucial to develop evidence-based practices for PD diagnosis, avoid pitfalls, assess impairment progression, and evaluate the effectiveness of treatment.

The main locus of PD is the dysfunction of the basal ganglia due to the degeneration of dopaminergic neurons in the substantia nigra. According to the model of Goberman & Coelho (2002), the decrease in dopamine leads to physiological changes at the laryngeal and respiratory levels and alters the timing of the lip/tongue movement initiation. These changes may lead to atypical prosody measures in PwPD regarding F0, intensity, speech rate and pause time. However, the model does not provide details of the underlying neural mechanisms and computations giving rise to these modifications. An influential neurocomputational model of speech production is the DIVA (Directions Into Velocities of Articulators)/GODIVA (Gradient Order DIVA) model of speech motor control (Tourville & Guenther, 2011; Bohland et al., 2010). According to DIVA/GODIVA, two cortico-basal ganglia loops are involved in the production of speech: the planning loop, which is responsible for buffering the phonological sequence to be produced, and the motor loop, which is in charge of generating the motor commands for producing the current phonological unit. Within these loops, pre-supplementary and supplementary motor areas are involved in the sequential and temporal aspects of speech production, the former representing the global structure of speech to be produced including syllable order and frame structure, and the latter representing individual articulatory gestures and their transitions within a speech sequence. Together with these two areas, the basal ganglia initiates speech sequences and activates automatic motor programs. Owing to basal ganglia dysfunction in PD, this automaticity is disturbed, thus leading to speech impairment (Guenther, 2016, pp. 279 - 283). The role of the basal ganglia in the production of speech sequences has also been suggested in other models of speech production (e.g., Kotz & Schwartze, 2016).

While these models of speech production address the temporal aspect of speech, and timing is crucial to convey prosody, their focus is not on prosody specifically. This may be due to the lack of data on the neurobiological bases of linguistic prosody. Compared to studies on 'emotional prosody' (Grandjean, 2020), which involves the acoustic encoding of the

expression of emotions and feelings, fewer studies have investigated the relationship between linguistic prosody and brain function (e.g., areas involved in using prosody to indicate a question or emphasis on a word, Aziz-Zadeh et al., 2010). However, a few regions have been identified as playing a potential role in linguistic prosody. These include the inferior frontal gyrus, ventral pre-motor cortex, supplementary and pre-supplementary motor areas and anterior cingulate cortex (for a review, see Guenther, 2016). Regarding the pure motor control of prosodic cues, which primarily involves movements of the laryngeal and respiratory muscles, the dorsal and ventral larynx motor cortex, cerebellum and supplementary motor area are believed to be involved (Brown et al., 2009; Dichter et al., 2018). However, the role of the basal ganglia in the generation and motor control of prosodic cues remains to be further investigated (Viñas-Guasch & Wu, 2017).

The main goal of the current study was to distinguish between the acoustic features of linguistic prosody that are reliably impaired in PD and those that remain unaltered. To achieve this goal, a systematic review and Bayesian meta-analysis were performed. Bayesian statistical modeling is particularly useful in our case since computing posterior distributions makes it possible to quantify evidence in favor of potential impairments beyond "binary" significance testing. Other advantages are the easy regularization (via conservative priors to reduce false positives), the easy implementation of robust techniques to reduce the impact of outliers and the ability to quantify the strength of null results (for a comprehensive review, see Wagenmakers et al., 2018). While the study's primary objective was to explore the impact of PD on acoustic features, it also aimed to explore the role of the basal ganglia in prosody production, using PD as an indication of basal ganglia dysfunction.

Methods

Search strategy

A systematic literature search was conducted on PubMed, PsycINFO and Web of Science. The initial aim was to perform a meta-analysis on the impact of PD (as a model of basal ganglia degeneration) and spinocerebellar ataxia (SCA, as a model of cerebellum degeneration) on the acoustic features of prosody. The following research terms were used: (prosody OR intonation OR inflection OR intensity OR pitch OR fundamental frequency OR speech rate OR rhythm) AND (Parkinson's disease OR spinocerebellar ataxia) AND (acoustic).

Two filters were also applied to search for articles in English published after 2000. Although the initial goal was to investigate both PD and SCA, the quantification of SCA was not feasible owing to the limited number of articles (see Study selection).

Study selection

Figure 1 shows the details of the selection procedure as a PRISMA flowchart (Haddaway et al., 2022). The selection was carried out by the second author (D.B.) between June 2022 and October 2022 and double-checked by the first author (J.F.). A total of 648 records were identified (Pubmed: 282, PsychInfo: 163, Web of Science: 203). Two hundred and eighty duplicate records were removed. A study was included if: (1) it was written in English; (2) it was published after 2000; (3) it reported original data for PwPD or PwSCA as well as controls; and (4) it included sufficient quantification of a speech task and the acoustic features of interest. Studies had to report at least one of the following features: F0, F0 variability, intensity, intensity variability, speech rate, articulation rate, pause duration and pause ratio (i.e., proportion of pauses during speaking). Both speech rate and articulation rate were included to provide a comprehensive understanding of the phenomenon of speaking rate. The former includes pause intervals while the latter excludes them. Regarding speech tasks, only studies that involved word and sentence production in a habitual manner were included, such as reading and spontaneous speech tasks. For instance, data on sustained /a/ production or diadochokinetic tasks were not included. Although studies measuring sustained phonation did report some of the features of interest (e.g., F0 and its variability and rate, respectively), they were not produced in the context of communicative utterances and were thus considered to be beyond the scope of the study. Based on the four above-mentioned criteria, 222 articles were excluded during screening. The full-text articles of the remaining 146 records were then read. Another 60 articles were excluded when checking for eligibility in terms of speech tasks and acoustic features. Only a few articles reported data for SCA. Owing to limited data samples for SCA and to ensure reliable effect sizes, a decision was made to focus on PD and not to include studies on SCA. A study was thus included if it reported data for PwPD as well as controls. Other articles were considered for inclusion based on the following additional criteria: (5) PwPD had to be without deep brain stimulation (DBS) and in an ON medication state (if there were medicated); (6) the studies had to report data using

linguistic prosody tasks (and not emotional prosody tasks); and (7) the studies had to report data that was not used in other studies included in the review. Some articles included data from participants who were also included in other articles with larger samples. To avoid any bias, the former articles that reported a subset of participants were excluded. A total of 110 articles were excluded in the second step of eligibility assessment (see Figure 1). The current meta-analysis included a total of 36 articles. Supplementary Material 1 provides the characteristics of PwPD including mean age and mean disease duration. A record was made of the excluded articles and of the reasons for their exclusion.

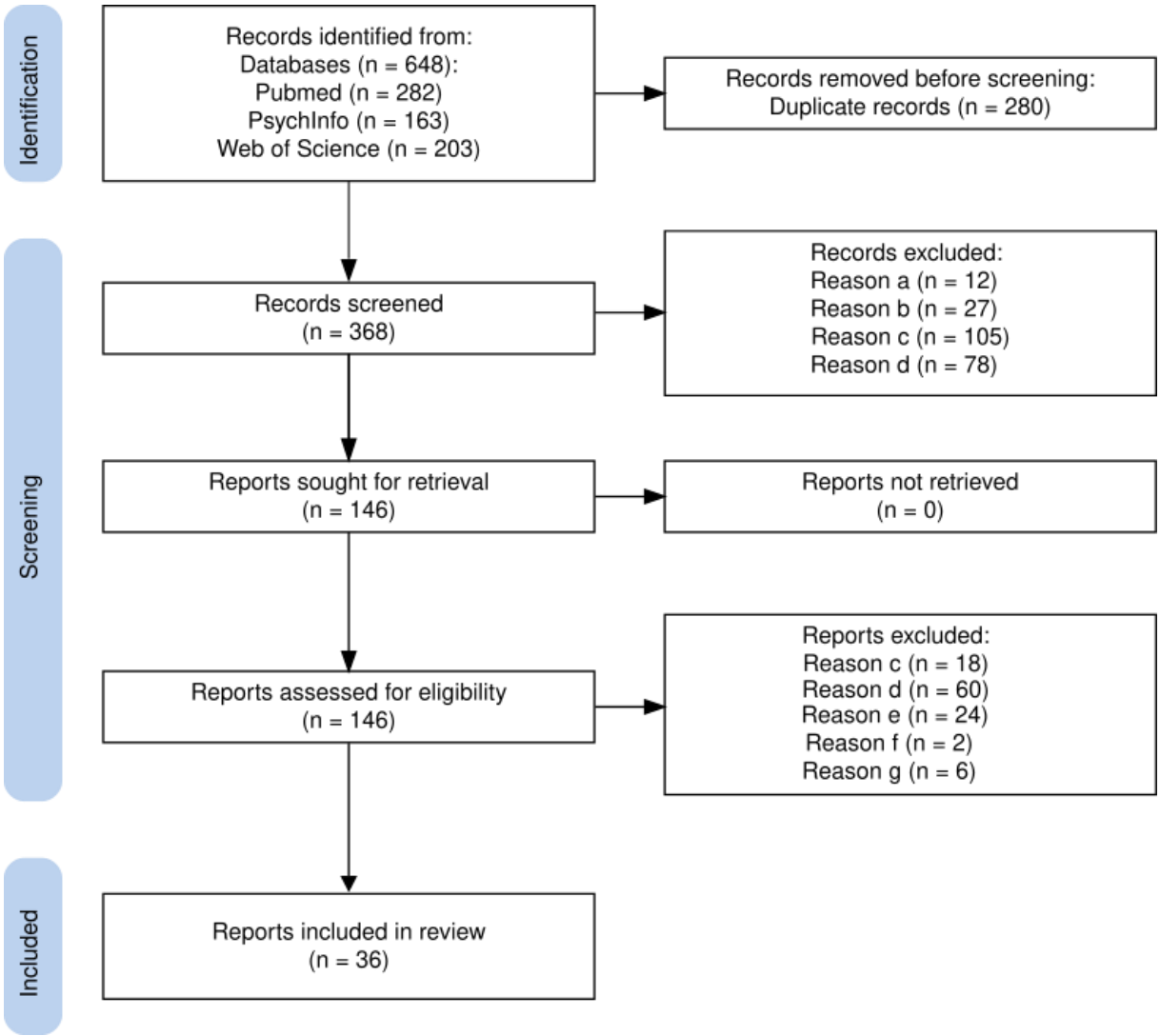

Figure 1: PRISMA flowchart. Reasons for exclusion: (1) not in English, (2) before 2000, (3) no original data on PD and controls, (4) no relevant speech task/acoustic feature, (5) Presence of DBS or medication removal, (6) emotional prosody, (7) duplicated data.

Data extraction

The following acoustic features were extracted: F0, F0 variability, intensity, intensity variability, speech rate, articulation rate, pause duration and pause ratio. When raw data were not directly available in the paper, they were requested from the authors or extracted using the Webplotdigitizer software (Rohatgi, 2022). Measures related to F0 are known to be sensitive to gender. Therefore, if an article reported measures for females and males as well as for the whole group, separate measures for females and males for F0 and its variability were included. In these cases, data on the whole group were not included. This decision was made to ensure that the same individuals were not included twice for the same task and for the same acoustic feature, once as females/males and once in the whole group. For other acoustic measures of interest, data on the whole group were included. All measures of variability were also extracted. However, if a study reported different measures for the same task, only standard deviation was included as this measure is less sensitive to extreme values than range, therefore providing a more reliable measure of F0 variability. Some of the studies reported mean pause duration while others reported median pause duration. In the present analyses, these two measures were grouped as they both reflect the central tendency of duration. In addition to acoustic features, the following data were extracted: number of participants in each group, mean age of participants, disease duration, treatment, language and details about the production task as well as the linguistic material used in each study. Data extraction was carried out by the second author (D.B.) between June 2022 and October 2022 and double-checked by the first author (J.F.) between December 2022 and February 2023.

Effect sizes were calculated using Hedges's *g* per feature and task for each study. In the current investigation, *g* represents the standardized mean difference between PwPD and controls, i.e., g = (meanPwPD – meanCtrl) / s.d.pooled , where s.d.pooled is the square root of (nPwPD- 1) s.d.PwPD + (nCtrl - 1) s.d.Ctrl) / (nPwPD + nCtrl – 2). Larger absolute values thus represent larger differences between the two groups. A negative effect size indicates that the value of interest was smaller for PwPD than for controls, and vice-versa. To calculate *g*, the *escalc* function (parameters: "SMD" and "vtype=AV") of the R package *metafor* (Viechtbauer, 2010) was used. If an article reported multiple measures for a feature (e.g., F0 variability collected during a reading task and during a run speech task), the effect size was calculated

separately for each task. In three articles, medians and quartiles were reported instead of means and standard deviations. For these studies, means and standard deviations were estimated using the R package *estmeansd* which implements the method of McGrath et al. (2020). Publication biases were assessed by checking for asymmetries in the funnel plots. The plots were inspected visually and analyzed to determine whether the observed effect sizes and the corresponding sampling variances were correlated using a rank correlation test. A high correlation suggests that the funnel plot is asymmetric, which may be a symptom of a publication bias (Begg & Mazumdar, 1994). Correlations were checked using the *ranktest* function of the R package *metafor* (Viechtbauer, 2010).

Hierarchical Bayesian model

Effect sizes were analyzed using hierarchical Bayesian regression models. An intercept-only model was built for each acoustic parameter of interest. To capture the impact of PD on prosody, the focus was directed on the variability of F0 and intensity, not on their absolute values (which are not linked to prosody functions *per se*). For completeness, the results regarding these features are available in Supplementary Material 2. Six models were built to examine the pooled effect size for the following parameters: F0 variability, intensity variability, speech rate, articulation rate, pause duration and pause ratio. More weight was attributed to effect sizes which were extracted from the studies with higher precision (i.e. fewer sampling errors) (Hedges & Olkin, 2014). As the effect sizes were nested within articles and the articles within investigated languages, a nested random-effect term was used to take into account the dependency between data in the same language/article as well as between-language/article heterogeneity. To reduce the influence of outliers, as for instance in Cox et al. (2023), a Student's t likelihood was used to model the output instead of the normal distribution that is classically used, as it accommodates extreme effect sizes in the tail of the distribution with minimal impact on mean and standard deviation. Weekly informative priors were defined. A normal distribution centered at 0 (no effect) with a standard deviation of 0.5 for the pooled effect size was used (thus discounting the possibility of effect sizes below -1 and above +1), as well as a positive truncated normal distribution centered at 0 with a standard deviation of 0.5 for heterogeneity. Regarding the degree of freedom, a gamma distribution was used with a shape parameter of 2 and a scale parameter of 0.1. For each

acoustic measure, the posterior distribution of the pooled effect size and heterogeneity were determined using the R package *brms* (Bürkner, 2017).

Based on posterior distributions, the estimated effect size for each model are reported. These values are interpreted in light of the corresponding 95% credible interval (CrI), evidence ratio and credibility score, to take into account the uncertainty and confidence associated with each measure. Ninety-five percent CrI is the interval within which there is a 95% probability that the true value of the effect size is included, given that the assumptions of the model are met. The evidence ratio quantifies the ratio of likelihood in favor of a given hypotheses (e.g., less F0 variability in PwPD than in controls) compared to other alternatives. Therefore, an evidence ratio K for a given hypothesis indicates that this hypothesis is K times more likely than the alternatives. Finally, the credibility score refers to the percentage of posterior samples in the direction of the hypothesis: for example, an effect size to be smaller than 0 for F0 variability (i.e., less variability in PwPD than in controls).

Results

Table 1 provides a summary of the results. The subsequent sections present the findings related to each feature of interest and evaluate the robustness of the results to potential publication biases. For each feature, a forest plot illustrates the distribution of estimated effect sizes (Figures 2 to 7). To facilitate comparison, consistent color coding is used for articles sharing the same language. A summary of results shown in each forest plot, including the position of the 95% CrI in relation to the null effect, is also provided in a Table in Supplementary Material 3.

Acoustic feature	No. of articles	No. of effect sizes	Estimated effect size	Evidence ratio	Credibility
F0 variability	20	47	-0.75 (-1, -0.48)	Inf	100%
Intensity variability	12	21	-0.31 (-0.87, 0.23)	8.73	90%
Speech rate	10	18	-0.35 (-1.11, 0.49)	6.27	86%
Articulation rate	14	24	-0.25 (-1.02, 0.48)	3.53	78%
Pause duration	9	20	0.5 (0.03, 1.02)	43.44	98%
Pause ratio	5	11	0.39 (-1.14, 2.09)	2.93	75%

Note. Numbers inside brackets are 95% credible intervals. F0 = fundamental frequency; Inf = infinite.

Table 1: Results for comparison of PwPD and controls for acoustic features of interest. Numbers between brackets are 95% CrIs.

Variability of F0 and intensity

F0 variability was the most common prosodic aspect reported in the studies. The analyses included 47 effect sizes from 20 articles for a total of 428 PwPD and 363 controls. The Bayesian hierarchical model revealed an overall estimated effect size of -0.75 with a 95% CrI of (-1, -0.48) as well as an overall estimated heterogeneity of 0.35 with a 95% CrI of (0.18, 0.58). Evidence ratio and credibility scores for estimated effect size were Inf and 100%, respectively. This means that all posterior samples for mean F0 variability in PwPD were smaller than all the posterior samples for the same measure in controls. 'Inf' means that all posterior samples were in the direction of the hypothesis. Figure 2 shows that the estimated effect sizes of four articles included the null in their CrIs. Except for these articles, all CrIs were totally distributed on the negative scales. Supplementary Figure 4.1 presents a funnel plot for variability of F0. The data did not reveal any likely publication bias (Kendall's tau = -0.01, p = 0.9). As discussed in the Methodology section, when available, gender-based data on F0 variability were included in the meta-analysis. To check for a potential effect of gender on the results, a few unplanned analyses were conducted. They involved examining potential differences between the results of the studies with gender-based data and that of the studies with group data. Additionally, the intercept model was compared to a model that included gender as a fixed-effect factor. The results did not show any effect related to gender (Supplementary Material 5).

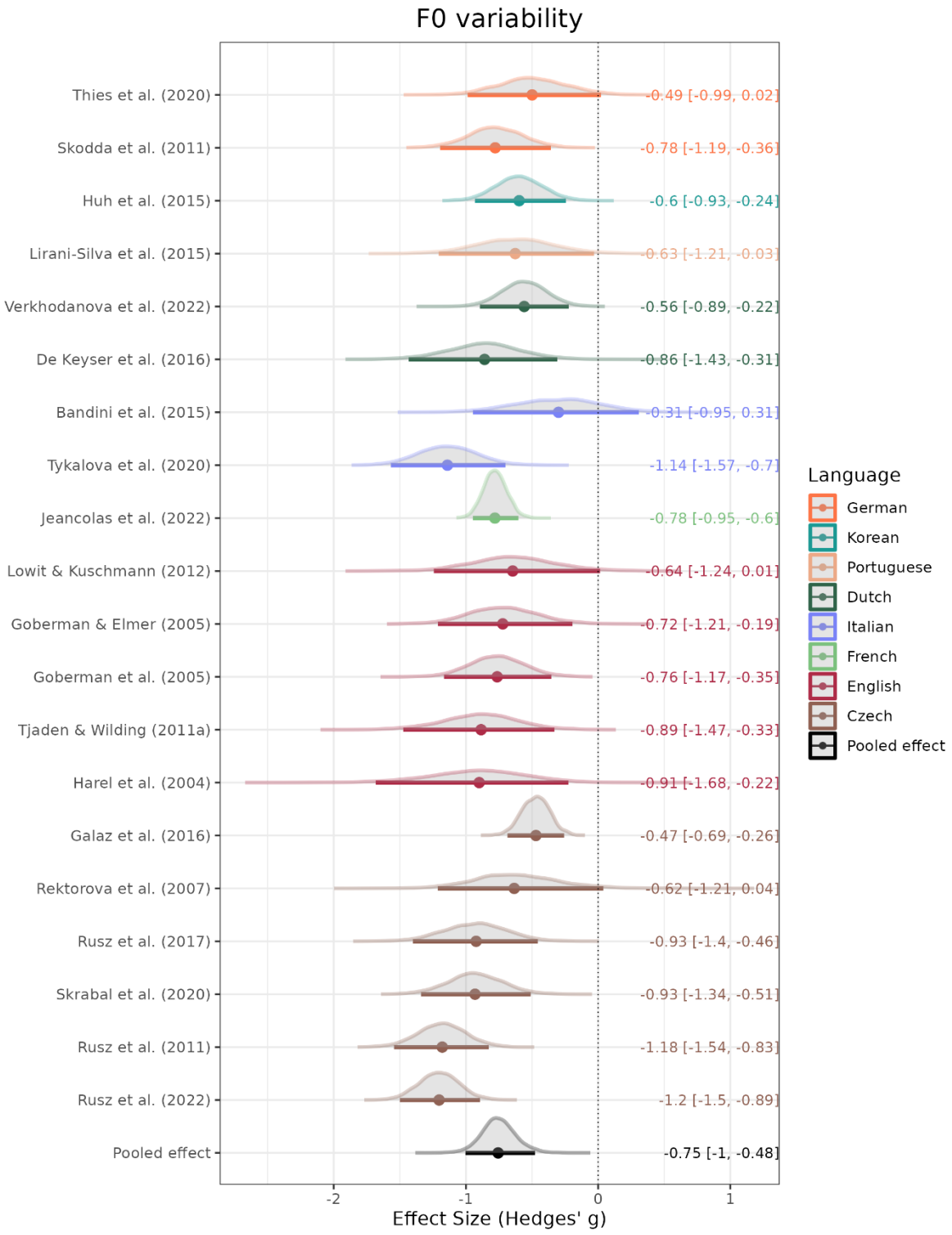

Figure 2: Forest plot for F0 variability estimates according to article. Shaded areas represent posterior probability density for each estimate. Numbers on right provide estimated Hedges' g and 95% CrI. Results for pooled effect are displayed at bottom of plot, providing a comprehensive summary of the findings.

Regarding intensity variability, the analyses included 21 effect sizes from 12 articles for a total of 322 PwPD and 247 controls. The Bayesian hierarchical model of intensity variability revealed an overall estimated effect size of -0.31 with a 95% CrI of (-0.87, 0.23), as well as an overall estimated heterogeneity of 0.53 with a 95% CrI of (0.27, 0.95). The evidence ratio and the credibility score for the estimated effect size were 8.73 and 90%, respectively. The forest plot in Figure 3 shows that nine articles included the null in their CrIs. The estimated effect sizes of the remaining four articles exhibited CrIs on the negative scales. Supplementary Figure 4.2 shows a funnel plot for variability of intensity. The data did not reveal any likely publication bias (Kendall's tau = -0.07, p = 0.67).

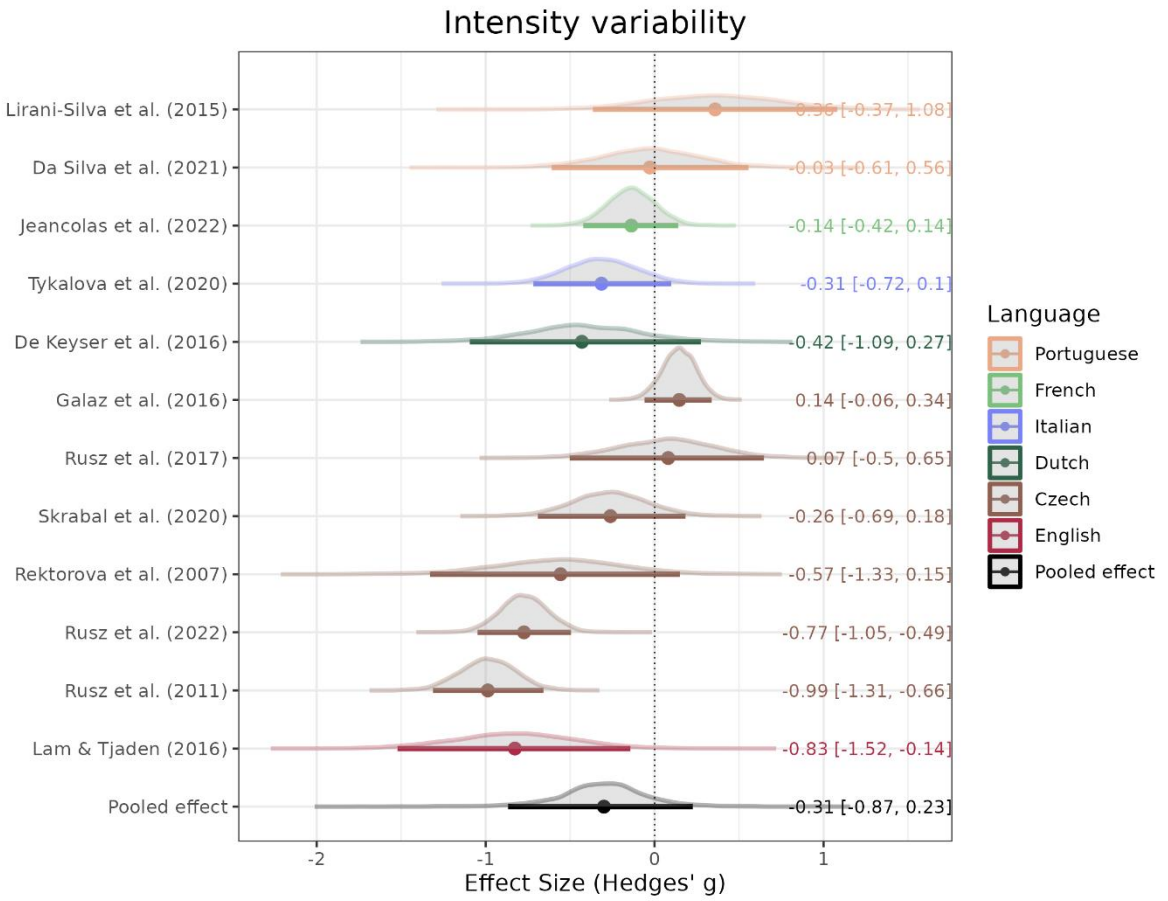

Figure 3: Forest plot for intensity variability estimates according to article. Shaded areas represent posterior probability density for each estimate. Numbers on right provide estimated Hedges' g and 95% CrI. Results for pooled effect are displayed at bottom of plot, providing a comprehensive summary of the findings.

Speaking rate

Regarding speech rate, our analyses were based on 18 effect sizes from 10 articles for a total of 196 PwPD and 168 controls. The Bayesian hierarchical model of speech rate showed an overall estimated effect size of -0.35 with a 95% CrI of (-1.11, 0.49), as well as an overall estimated heterogeneity of 0.49 with a 95% CrI of (0.02, 0.77). Evidence ratio and credibility score for estimated effect size were 6.27 and 86%, respectively. The forest plot in Figure 4 shows that the estimated effect size of six articles included the null in their CrIs. The estimated effect sizes of the remaining three articles exhibited CrIs totally on the negative scales. Supplementary Figure 4.3 presents a funnel plot for speech rate. The analysis revealed a potential publication bias (Kendall's tau = -0.39, p = 0.03).

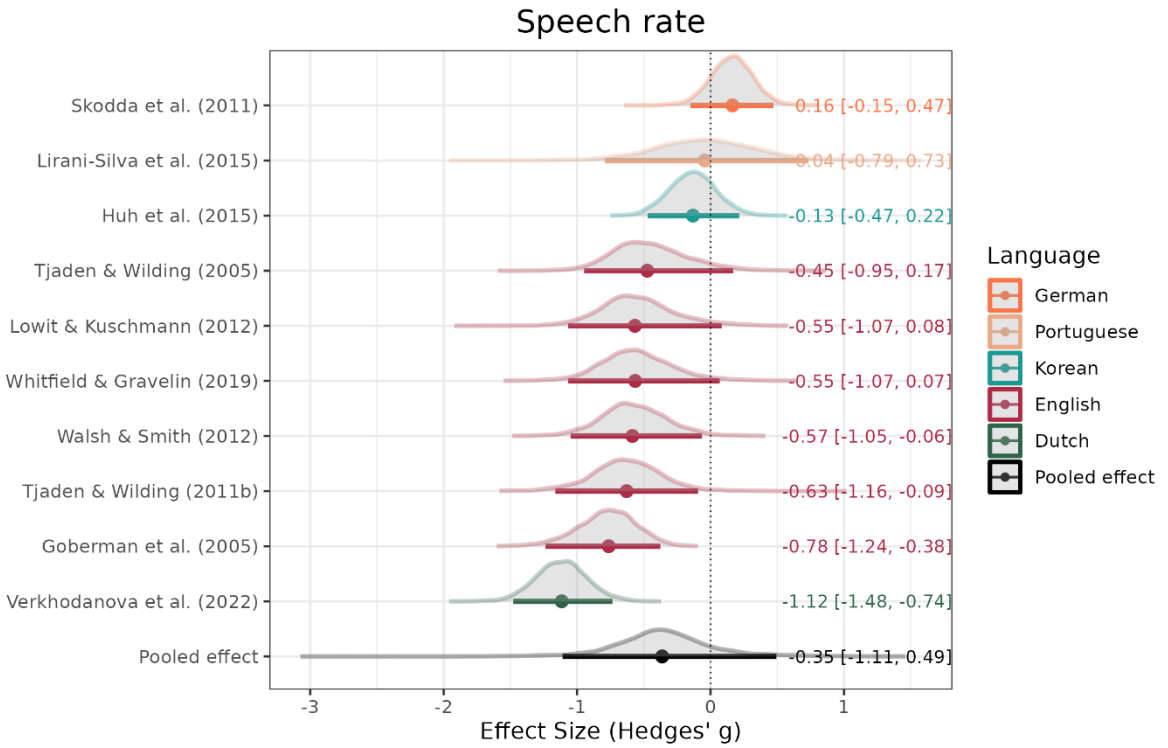

Figure 4: Forest plot for speech rate estimates according to article. Shaded areas represent posterior probability density for each estimate. Numbers on the right provide estimated Hedges' g and 95% CrI. Results for pooled effect are displayed at bottom of plot, providing a comprehensive summary of the findings.

Regarding articulation rate, our analyses were based on 24 effect sizes from 14 articles for a total of 294 PwPD and 255 controls. The Bayesian hierarchical model of showed an overall estimated effect size of -0.25 with a 95% CrI of (-1.02, 0.48), as well as an overall estimated heterogeneity of 0.39 with a 95% CrI of (0.03, 0.89). Evidence ratio and credibility score for estimated effect size were 3.53 and 78%, respectively. As shown in Figure 5, the estimated effect sizes for only two articles does not include the null in their CrIs. The CrIs of these two articles were distributed on the negative scales. Supplementary Figure 4.4 shows a funnel plot for articulation rate. The data did not reveal any likely publication bias (Kendall's tau = -0.03, p = 0.82).

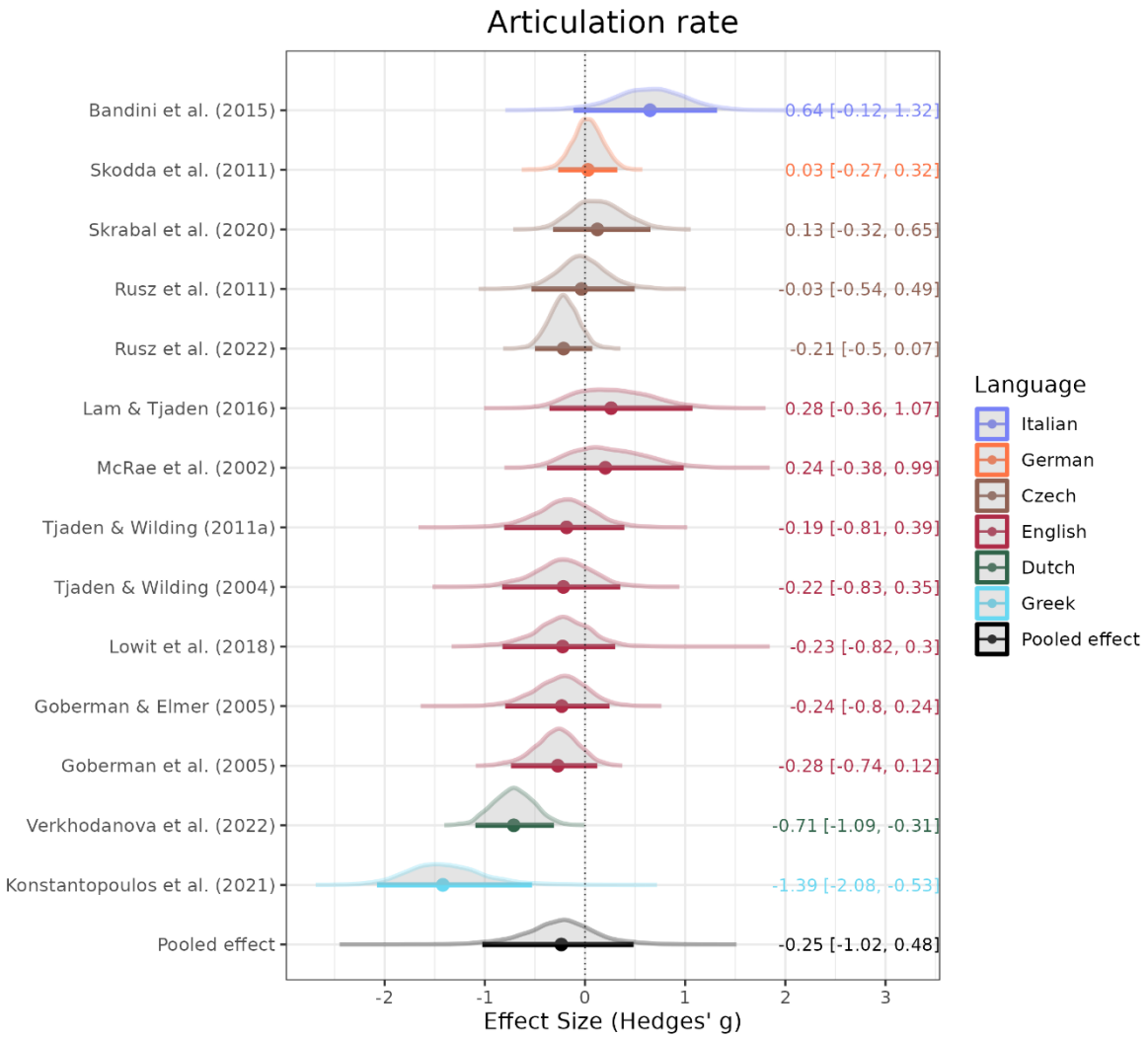

Figure 5: Forest plot for articulation rate estimates according to article. Shaded areas represent posterior probability density for each estimate. Numbers on right provide estimated

Hedges' g and 95% CrI. Results for pooled effect are displayed at bottom of plot, providing a comprehensive summary of the findings.

Pause

Regarding pause duration, 20 effect sizes were extracted from nine articles for a total of 196 PwPD and 223 controls. The Bayesian hierarchical model of pause duration showed an overall estimated effect size of 0.5 with a 95% CrI of (0.03, 1.02), as well as an overall estimated heterogeneity of 0.3 with a 95% CrI of (0.05, 0.74). Evidence ratio and credibility score for estimated effect size were 43.44 and 98%, respectively. This implies that approximately 98% of PwPD samples would show a longer pause duration than that of the samples of control participants. Figure 6 shows that the effect size estimates for two articles included the null in their CrIs. All other articles exhibited CrIs totally distributed on the positive scales. Supplementary Figure 4.5 shows a funnel plot for pause duration. The data did not reveal any likely publication bias (Kendall's tau = 0.09, p = 0.6).

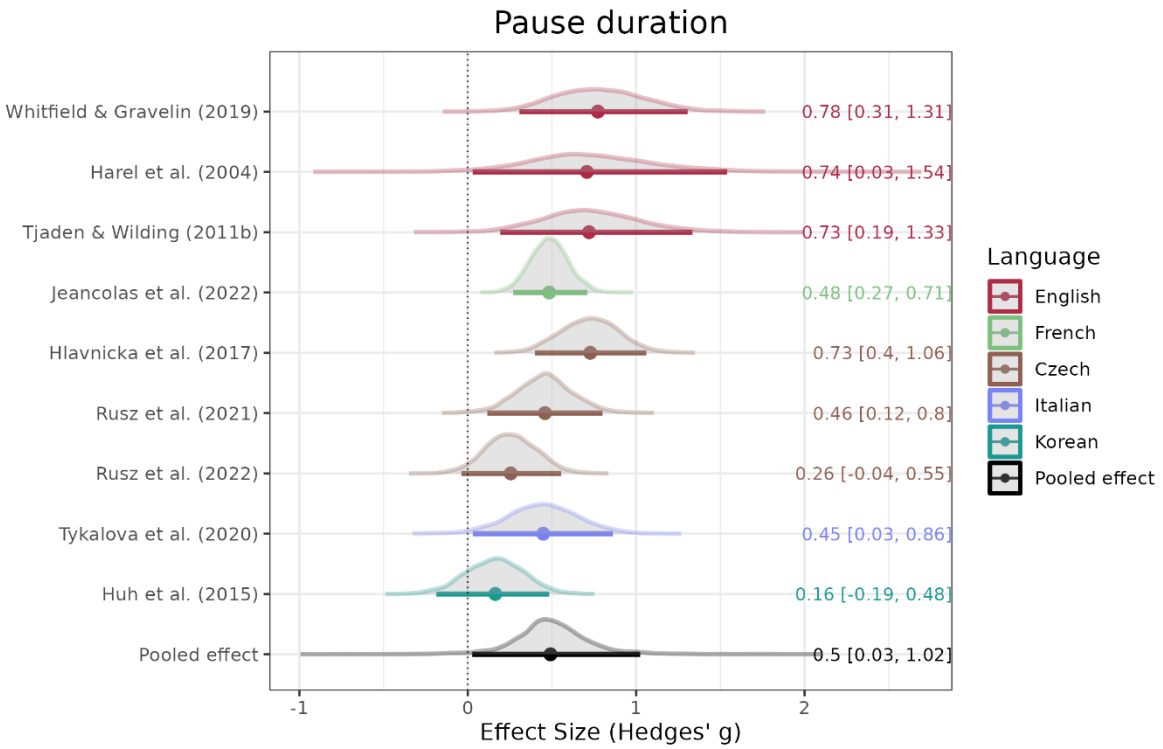

Figure 6: Forest plot for pause duration estimates according to article. Shaded areas represent posterior probability density for each estimate. Numbers on right provide estimated Hedges' g and 95% CrI. Results for pooled effect are displayed at bottom of plot, providing a comprehensive summary of the findings.

The analysis of pause ratio was based on 11 reported effect sizes from five articles for a total of 127 PwPD and 87 controls. The Bayesian hierarchical model of pause ratio showed an overall estimated effect size of 0.39 with a 95% CrI of (-1.14, 2.09), as well as an overall estimated heterogeneity of 0.71 with a 95% CrI of (0.04, 2.05). Evidence ratio and credibility score for estimated effect size were 2.93 and 75%, respectively. Figure 7 shows that the estimated effect sizes of three articles included the null in their CrIs. The CrIs for the remaining two articles were totally distributed on the positive scales. Supplementary Figure 4.6 shows a funnel plot for pause ratio. The data on pause ratio revealed a likely publication bias (Kendall's tau = 0.51, p = 0.04).

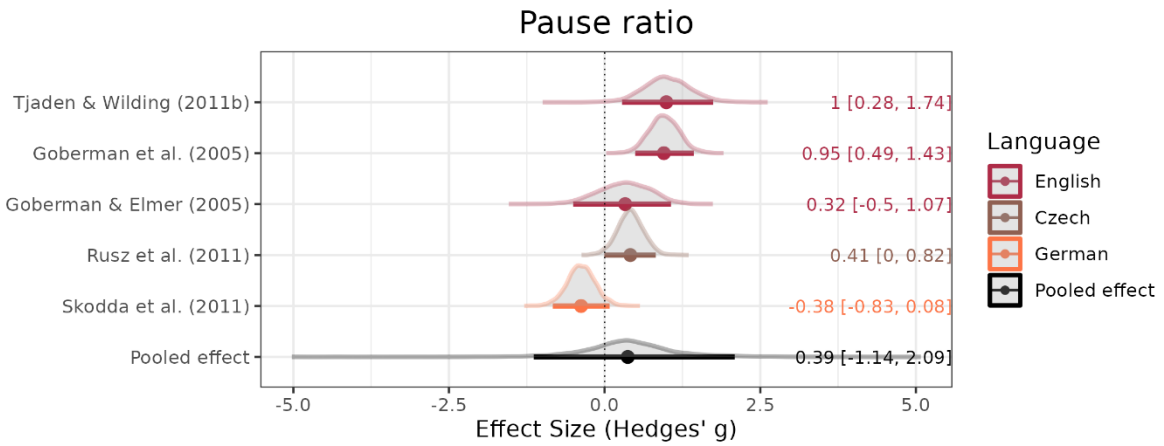

Figure 7: Forest plot for pause ratio estimates according to article. Shaded areas represent posterior probability density for each estimate. Numbers on right provide estimated Hedges' g and 95% CrI. Results for pooled effect are displayed at bottom of plot, providing a comprehensive summary of the findings.

Discussion

Overview of the findings

Prosody is one of the speech and language dimensions most often reported as being impaired in PwPD. The primary objective of this meta-analysis was to systematically investigate the effect of PD on the acoustic features of linguistic prosody. It also aimed to shed light on the role of the basal ganglia in prosody production, using PD as an example of basal ganglia dysfunction. F0 and intensity variability, speech and articulation rate, as well as pause duration and ratio were examined. Overall, strong evidence was found in favor of the impact of PD on some acoustic features related to prosody. The analysis of potential publication biases supported the reliability of these findings, as only the data on pause ratio and speech rate indicated a likely publication bias.

The strongest and most reliable effect was found to be that of F0 variability, suggesting that reduced pitch variation is characteristic of the speech of PwPD. This is consistent with the results of a recent study in a large sample of French PwPD (Vaysse, Ghio, et al., 2022). The modulation of F0 is crucial to convey various prosodic cues such as intonation, tone and phrasing (Arvaniti, 2020). Future studies should investigate whether this alternation at the phonetic level also impacts the production of cues at the phonological level, which are

expressed through F0. To date, only a few studies have adopted such approaches to investigate the phonology of prosody in PD (Frota et al., 2021; Lowit & Kuschmann, 2012) and the results are not conclusive. To enhance understanding of the impact of PD on prosodic cues signaled through F0, investigations into tonal languages would be particularly interesting. In tonal languages, F0 is involved in the lexical realization. While the latter have received less attention in PD literature than non-tonal languages, perceptual evidence from Cantonese and Mandarin speakers suggests that the production of tones is relatively robust in such individuals (Whitehill et al., 2010). It has been reported that PD affects the control of tone and intonation differently. Intonation involves F0 modulation which spans utterances and was reported to improve after a speech therapy intervention, unlike tone (Whitehill et al., 2011). Further studies are needed to explore the potentially differential impact of PD on prosodic phenomena involving the variation of F0. To this end, the study of tonal languages would be beneficial.

The second strongest and most convincing effect was found to be that of pause duration, indicating that PwPD tend to produce longer pauses during speech production. Abnormal pause characteristics in terms of their duration, placement (at linguistic boundary or not) and frequency are thought to reflect various impairments in PD, including dysfunctions of speech planning, speech motor control and lexical retrieval (e.g., Ahn et al., 2014; Andrade et al., 2023; Lee et al., 2019). “Inappropriate silences” is one of the perceptual descriptions of hypokinetic dysarthria as defined by Darley et al., (1969b), which is commonly used in clinical settings to classify different types of dysarthria. It refers to instances in speech when “there are inappropriate silent intervals” (Darley et al., 1969b, p. 268), and is thought to be linked with difficulties in initiating movements (Darley et al., 1969a). This description may thus refer to pause duration issues but also other phenomena such as inappropriate placement of silences and abnormal pause frequency. Importantly, pause duration can be defined in various ways. It may be measured within or between utterances or be categorized as short versus long time intervals. This study included all types of pause duration and showed that the duration of pauses is generally longer in PwPD than in controls. Further studies should investigate whether the production of different types of pauses is impacted differently in PD. In the present study, pause ratio was also examined. Although a positive estimated effect size was found, the 95% CrI spans widely across both positive and negative values (see also Table 1 for evidence ratio and credibility score). Moreover, this analysis was potentially affected by a

publication bias. Several reasons may lead to publication bias in meta-analyses (Thornton & Lee, 2000). Regarding pause ratio in the current meta-analysis, several issues could be mentioned: very few studies were available for this analysis (see Table 1) with small studies tending to lead to greater positive effect sizes than larger studies. In addition, small studies with null results were missing (see Supplementary Figure 4.6). Given these factors, persuasive evidence of an effect of PD on pause ratio cannot be provided.

The impact of PD on pause duration can be understood in light of the evidence of role of the basal ganglia in speech production. In a study comparing PwPD with and without DBS, Ahn et al., (2014) observed that DBS modulates pause characteristics as stimulation led to more frequent and shorter pauses, as well as a greater proportion of pauses in non-linguistic boundaries. This highlights the role of brain circuits in the production of pauses, including the basal ganglia. Although further studies are needed for confirmation, these findings may be interpreted within the framework of DIVA/GODIVA (Guenther, 2016) discussed in the Introduction. The disturbance of the circuit responsible for speech timing and sequencing, involving the basal ganglia, the pre-supplementary and supplementary motor areas, may lead to abnormal pause characteristics. In addition to issues related to speech motor control, abnormal pauses may reflect language deficits such as difficulties in lexical and syntactic processing (Lee et al., 2019; Mack et al., 2015). Importantly, basal ganglia dysfunction is thought to lead to word production difficulties (Camerino et al., 2022). According to Bohsali & Crosson (2016), two different cortical-subcortical loops, both involving the basal ganglia, underlie lexical selection and the activation of appropriate phonological and articulatory representations. The latter loop, but not the former, involves the left inferior frontal gyrus. However, the way these various levels of processing are affected remains to be defined. Another research avenue is to investigate the role of the basal ganglia in relation to executive functions, such as executive control, working memory and attention, known to improve the efficiency and fluidity of language (Jacquemot & Bachoud-Lévi, 2021). Finally, beyond speech and language, pause duration has recently been suggested to be a crucial feature in PD, as it has been related to cognitive function (Andrade et al., 2023). If this relationship is confirmed by future studies, pause duration could potentially be a valuable speech marker of cognitive decline in PD.

One of the speech aspects frequently reported as altered in PD is reduced loudness. The Lee Silverman Voice Treatment (LSVT) LOUD method, a well-known speech therapy intervention method, targets loudness during treatment in PwPD (Ramig et al., 2018). Importantly, intensity is one of the important phonetic parameters used to express prosody. For instance, stress is mostly expressed using a combination of intensity and duration changes during speech production (Arvaniti, 2020). Controlling intensity is therefore important for the production of prosodic cues. A negative estimated effect size for intensity variability was found, indicating that PwPD modulate the intensity of their speech less than controls. However, the evidence supporting the effect was not robust. This is consistent with the observation that intensity variability may differ between hyperkinetic and hypokinetic dysarthria, but not when compared to controls (Rusz, Tykalova, et al., 2021). In their investigation on lexical stress production in PD, Exner et al. (2023) observed no effect of PD on intensity variability, measured using the pairwise variability index, which is consistent with the present findings. Overall, intensity variability does not seem to constitute a prosody impairment in PD.

Speaking rate is another aspect that is commonly acknowledged as being impaired in PD. Unlike aspects such as F0 variability on which there is a consensus regarding the direction of the effect, different patterns have been reported for speaking rate, including too slow, too fast, and variable (Blanchet & Snyder, 2009; Moreau & Pinto, 2019). In this study, both speech rate and articulation rate were examined. A negative estimated effect size for speech rate was found, indicating that the latter is reduced in PD. This is consistent with the fact that PD impacts speech timing. However, the strength and credibility of the effect was similar to the effect found for intensity variability. It was thus less robust than the effect for F0 variability and pause duration. Furthermore, the analysis on speech rate may have been influenced by a potential publication bias, thus caution should be applied when interpreting these results. Regarding the effect for articulation rate, although the estimated effect size was negative, no evidence supports the effect for articulation rate given its founded credibility. Speech rate thus seems to be a better indicator of potentially altered speaking rate in PD than articulation rate. This is consistent with our results regarding pause duration and suggests that PD has a greater impact on factors that contribute to pausing, such as planning or breathing, than on articulation itself.

Speech rate is one of the dimensions used in the therapeutic management of speech and language impairments in PD. In these therapies, PwPD are asked to slow their speech rate during exercises, regardless of their usual rate (Blanchet & Snyder, 2009; Martens et al., 2015). However, the effectiveness of such exercises in terms of intelligibility improvement is debated. Knowles et al., (2021) explored changes in intelligibility across a wide range of self-selected speech rates from very slow to very fast. They studied PwPD with medication, PwPD with DBS and controls, and observed that while fast rates were associated with decreased intelligibility, slowing one's speaking rate was not associated with intelligibility changes. However, PwPD with DBS benefited from slowing their speech rate. In addition, speech intelligibility was found to be poorer in the monologue task than in the sentence-reading task, with greater group differences observed in the monologue task. Further studies are needed to understand whether and how changes in the speaking rate in PwPD impact the perception of listeners, in terms of intelligibility and comprehension. To do so, the impact of speech tasks and the characteristics of PwPD should be investigated as they may be influential (Knowles et al., 2021; Lowit et al., 2018). Interestingly, Illner et al. (2022) observed that speech and articulation rates produced by de novo PwPD remained unchanged in monologues and reading tasks compared to controls. Furthermore, their findings suggest that potential changes in speaking rates may be dependent on PD progression. The aim of this study was to examine the reliability of various techniques to assess speech and articulation rates in PD, idiopathic rapid eye movement sleep behavior disorder, and multiple system atrophy compared to controls. Note that this study was not identified in our article search, possibly owing to the keywords used. This may raise concerns about the exhaustiveness of our procedure. However, our method and data are available for future research extensions. In addition, the findings of Illner et al. (2022) align with our conclusion and point to limited evidence of an impact of PD on speech and articulation rates, thus supporting the representativeness of our results.

In recent years, there has been an increasing use of automatic tools to monitor PD evolution and a growing body of research on speech biomarkers of Parkinson's disease (PD) (e.g., Kouba et al., 2022; Jeancolas et al., 2021; Laganaro et al., 2021; Rusz et al., 2021). Our results suggest that F0 variability and pause duration may be the most reliable features for this purpose. Thanks to advances in extraction algorithms for pathological speech (e.g., Vaysse, Astésano, et al., 2022), these features could easily be integrated into clinical settings. Promising approaches for the identification of vocal markers of certain pathologies are based

on machine learning methods and are thus data-driven, not relying on features informed by meta-analyses. However, the generalizability and specificity of such models is an issue (Hansen et al., 2023; Rybner et al., 2022). Future studies should compare the generalizability and specificity of features informed by meta-analyses with those based on machine learning models. To this aim, language-specific phenomena should be taken into account. The constraints which are specific to the language of PwPD may impact the nature and degree of speech and language impairments (Pinto et al., 2017). Similar observations were made in other clinical populations. Parola et al. (2023) and Fusaroli et al. (2022) found that reliable cross-linguistically acoustic profiles respectively of schizophrenia and autism are unlikely. To implement an evidence-based approach, cross-linguistic studies should thus be undertaken to understand the interplay between linguistic constraints and abnormal acoustic features in PD.

Altogether, the present results suggest that PD impacts the control of the main parameters that convey prosody, namely F0, duration and to a lesser extent, amplitude. Whether these parameters are controlled independently or not remains to be addressed (for a review, see Guenther (2016), pp. 252-256). In their study examining the influence of PD on the production of contrastive stress, Tykalova et al. (2014) used an aggregated measure to combine the effects of individual acoustic cues on the production of contrastive stress. Although PwPD could employ strategies similar to controls to signal contrastive stress, they demonstrated a reduced ability to modulate the acoustic cues. The authors suggested that PwPD might face increased challenges when needing to simultaneously modulate pitch, intensity, and duration. The fact that speech production models, including those dealing with motor control or higher levels of processing (e.g., Hickok, 2014; Pickering & Gambi, 2018; Tourville & Guenther, 2011), do not address prosody complexifies our understanding of this issue. The present findings may provide the beginning of an answer, as they suggest that the brain circuits controlling these parameters are at least partly shared and include the basal ganglia.

Meta-analyses on speech and language processes can provide robust evidence to inform and refine theoretical models. This may be achieved when meta-analyses examine specific claims of the models. For instance, Bürki et al. (2020) conducted a meta-analysis on semantic interference during word production when participants were asked to name pictures while ignoring semantically interfering words (e.g., dog-cat). The meta-analysis aimed to inform models which describe the process involved in selecting the appropriate

representation of a word during language production (i.e., lexical access). To do so, different factors suggested in psycholinguistic models were examined. Another example is the meta-analyses conducted by Meekings and Scott (2021) on error monitoring during speech production using neuroimaging data. The aim was to examine the role of the posterior superior temporal gyrus in error monitoring as developed in speech production models (Guenther & Hickok, 2016). The current meta-analysis was not specifically designed with this objective in mind. This should be the direction of future work and could be achieved by examining specific prosody phenomena including both behavioral and neuroimaging data. However, as mentioned in the introduction, most neuroimaging studies to date have focused on emotional prosody. Further studies on the neurobiological bases of linguistic prosody are needed to better understand the neurocognitive mechanisms underlying linguistic prosody impairments in PD and to refine speech production models in the light of such results.

Limitations

The current study has some limitations. First, the profile of PwPD and tasks included in the meta-analysis was not fully considered. Given the mixed evidence on the impact of DBS on speech production (Moreau & Pinto, 2019), PwPD who had undergone DBS and those who had their medication removed were excluded. These criteria helped reduce heterogeneity in our study, which is a positive outcome. However, some factors were not controlled for, such as disease duration and severity or speech and language therapy followed by PwPD. Moreover, the nature of the speech task may impact the production of prosodic cues as the demands are different during isolated word repetition, text reading or spontaneous speech. These issues may thus impact the conclusions and generalizability of the present study. In addition, the studies were not completely homogeneous regarding the methodology and measurement tools used in their acoustic analyses. Using standardized procedures for data collecting and analyzing (e.g., Rusz, Tykalova, et al., 2021) might help in reaching a more comprehensive understanding of how PD affects prosody and verbal communicative abilities in general. Another limitation concerns the languages included in this study. As shown in the forest plots in the Results section, the present meta-analysis examined a limited number of languages, some being more represented than others. Language-specific phenomena can thus be expected to impact the conclusions. Differences between tonal and non-tonal languages

may be expected (Whitehill., 2010; Pinto et al., 2017). Finally, by studying PD, one of the aims of this study was to investigate the role of the basal ganglia in the production of prosody. However, a more direct approach based on neuroimaging data is necessary to establish whether and how alterations in the basal ganglia itself may lead to impairments in prosody production.

Conclusion

A systematic review and a meta-analysis were conducted to evaluate the impact of PD on linguistic prosody production, as measured by acoustic analysis. Reliable evidence of reduced F0 modulation and longer pauses in PwPD than in controls was found. These two features are therefore the most reliable changes in PD related to prosody. Evidence of reduced intensity modulation and reduced speech rate in PwPD than in controls was also found. These findings are clinically useful and may guide the choice of PD management, including the parameters which could be used in the tools to monitor PD evolution. Finally, they suggest that PD impacts the control of the main parameters that convey prosody: F0, duration and, to a lesser degree, amplitude. Since the degeneration of the basal ganglia is the primary locus of PD, this suggests that the brain circuits controlling acoustic parameters of prosody are at least partly shared and include the basal ganglia. The current study focused on the phonetic realization of prosodic cues. Further studies should highlight whether and how PD impacts prosody at higher levels of processing.

Data availability

The data that support the findings of this study are available in Zenodo at http://doi.org/10.5281/zenodo.11072486.

Acknowledgments

We would like to thank Laurent Ott for his help with illustrations. This article was supported by the European center for humanities and social sciences (MESHS-Lille, France) and by the French National Agency for Research (project ANR-21-CE28-0011).

References

Ahn, J. S., Van Lancker Sidtis, D., & Sidtis, J. J. (2014). Effects of Deep Brain Stimulation on Pausing During Spontaneous Speech in Parkinson's Disease. *Journal of Medical Speech-Language Pathology*, *21*(3), 179–186.

Andrade, E. I. N., Manxhari, C., & Smith, K. M. (2023). Pausing before verb production is associated with mild cognitive impairment in Parkinson's disease. *Frontiers in Human Neuroscience*, *17*. https://www.frontiersin.org/articles/10.3389/fnhum.2023.1102024

Arora, S., Baghai-Ravary, L., & Tsanas, A. (2019). Developing a large scale population screening tool for the assessment of Parkinson's disease using telephone-quality voice. *The Journal of the Acoustical Society of America*, *145*(5), 2871-2884.

Arnold, C., Gehrig, J., Gispert, S., Seifried, C., & Kell, C. A. (2014). Pathomechanisms and compensatory efforts related to Parkinsonian speech. *NeuroImage: Clinical*, *4*(Supplement C), 82–97. https://doi.org/10.1016/j.nicl.2013.10.016

Arvaniti, A. (2020). The Phonetics of Prosody. In A. Arvaniti, *Oxford Research Encyclopedia of Linguistics*. Oxford University Press. https://doi.org/10.1093/acrefore/9780199384655.013.411

Atkinson-Clement, C., Sadat, J., & Pinto, S. (2015). Behavioral treatments for speech in Parkinson's disease: Meta-analyses and review of the literature. *Neurodegenerative Disease Management*, *5*(3), 233–248. https://doi.org/10.2217/nmt.15.16

Aziz-Zadeh, L., Sheng, T., & Gheytanchi, A. (2010). Common Premotor Regions for the Perception and Production of Prosody and Correlations with Empathy and Prosodic Ability. *PLOS ONE*, *5*(1), e8759. https://doi.org/10.1371/journal.pone.0008759

Bandini, A., Giovannelli, F., Orlandi, S., Barbagallo, S. D., Cincotta, M., Vanni, P., Chiaramonti, R., Borgheresi, A., Zaccara, G., & Manfredi, C. (2015). Automatic identification of dysprosody in idiopathic Parkinson's disease. *Biomedical Signal Processing and Control*, *17*, 47-54. https://doi.org/10.1016/j.bspc.2014.07.006

Basirat, A., Patin, C., & Moreau, C. (2018). Relationship between Perception and Production of Intonation of French in Parkinson's Disease. *In Proceedings of the 9th International*

Conference on Speech Prosody 2018, 809–813. https://doi.org/10.21437/SpeechProsody.2018-163

Blanchet, P. G., & Snyder, G. J. (2009). Speech rate deficits in individuals with Parkinson's disease: A review of the literature. *Journal of Medical Speech - Language Pathology*, *17*(1), 1–8.

Begg, C. B., & Mazumdar, M. (1994). Operating characteristics of a rank correlation test for publication bias. *Biometrics*, *50*(4), 1088–1101. https://doi.org/10.2307/2533446

Bohland, J. W., Bullock, D., & Guenther, F. H. (2010). Neural Representations and Mechanisms for the Performance of Simple Speech Sequences. *Journal of Cognitive Neuroscience*, *22*(7), 1504–1529. https://doi.org/10.1162/jocn.2009.21306

Bohsali, A., & Crosson, B. (2016). The Basal Ganglia and Language: A Tale of Two Loops. In J.-J. Soghomonian (Ed.), *The Basal Ganglia: Novel Perspectives on Motor and Cognitive Functions* (pp. 217–242). Springer International Publishing. https://doi.org/10.1007/978-3-319-42743-0_10

Brown, S., Laird, A. R., Pfordresher, P. Q., Thelen, S. M., Turkeltaub, P., & Liotti, M. (2009). The somatotopy of speech: Phonation and articulation in the human motor cortex. *Brain and Cognition*, *70*(1), 31–41. https://doi.org/10.1016/j.bandc.2008.12.006

Bürki, A., Elbuy, S., Madec, S., & Vasishth, S. (2020). What did we learn from forty years of research on semantic interference? A Bayesian meta-analysis. *Journal of Memory and Language*, *114*, 104125. https://doi.org/10.1016/j.jml.2020.104125

Bürkner, P.-C. (2017). brms: An R Package for Bayesian Multilevel Models Using Stan. *Journal of Statistical Software*, *80*, 1–28. https://doi.org/10.18637/jss.v080.i01

Camerino, I., Ferreira, J., Vonk, J. M., Kessels, R. P. C., de Leeuw, F.-E., Roelofs, A., Copland, D., & Piai, V. (2022). Systematic Review and Meta-Analyses of Word Production Abilities in Dysfunction of the Basal Ganglia: Stroke, Small Vessel Disease, Parkinson's Disease, and Huntington's Disease. *Neuropsychology Review*. https://doi.org/10.1007/s11065-022-09570-3

Cox, C., Bergmann, C., Fowler, E., Keren-Portnoy, T., Roepstorff, A., Bryant, G., & Fusaroli, R. (2023). A systematic review and Bayesian meta-analysis of the acoustic features of infant-

directed speech. *Nature Human Behaviour*, *7*(1), Article 1. https://doi.org/10.1038/s41562-022-01452-1

Darley, F. L., Aronson, A. E., & Brown, J. R. (1969a). Clusters of deviant speech dimensions in the dysarthrias. *Journal of speech and hearing research*, *12*(3), 462-496.

Darley, F. L., Aronson, A. E., & Brown, J. R. (1969b). Differential Diagnostic Patterns of Dysarthria. *Journal of Speech and Hearing Research*, *12*(2), 246–269. https://doi.org/10.1044/jshr.1202.246

Darling, M., & Huber, J. E. (2011). Changes to articulatory kinematics in response to loudness cues in individuals with Parkinson's disease. *Journal of Speech, Language, and Hearing Research*, *54*(5), 1247-1259. https://doi.org/10.1044/1092-4388(2011/10-0024)

da Silva, J. M. S., Gomes, A. de O. C., da Silva, H. J., de Vasconcelos, S. J., de Sales Coriolano, M. das G. W., & de Lira, Z. S. (2021). Effect of Resonance Tube Technique on Oropharyngeal Geometry and Voice in Individuals with Parkinson's Disease. *Journal of Voice*, *35*(5), 807.e25-807.e32. https://doi.org/10.1016/j.jvoice.2020.01.025

De Keyser, K., Santens, P., Bockstael, A., Botteldooren, D., Talsma, D., De Vos, S., Van Cauwenberghe, M., Verheugen, F., Corthals, P., & De Letter, M. (2016). The Relationship Between Speech Production and Speech Perception Deficits in Parkinson's Disease. *Journal of Speech, Language, and Hearing Research*, *59*(5), 915-931. https://doi.org/10.1044/2016_JSLHR-S-15-0197

Dichter, B. K., Breshears, J. D., Leonard, M. K., & Chang, E. F. (2018). The Control of Vocal Pitch in Human Laryngeal Motor Cortex. *Cell*, *174*(1), 21-31.e9. https://doi.org/10.1016/j.cell.2018.05.016

Exner, A. H., Francis, A. L., MacPherson, M. K., Darling-White, M., & Huber, J. E. (2023). The Effects of Speech Task on Lexical Stress in Parkinson's Disease. *American Journal of Speech-Language Pathology*, *32*(2), 506–522. https://doi.org/10.1044/2022_AJSLP-22-00185

Frota, S., Cruz, M., Cardoso, R., Guimarães, I., Ferreira, J. J., Pinto, S., & Vigário, M. (2021). (Dys)Prosody in Parkinson's Disease: Effects of Medication and Disease Duration on Intonation and Prosodic Phrasing. *Brain Sciences*, *11*(8), Article 8. https://doi.org/10.3390/brainsci11081100

Fusaroli, R., Grossman, R., Bilenberg, N., Cantio, C., Jepsen, J. R. M., & Weed, E. (2022). Toward a cumulative science of vocal markers of autism: A cross-linguistic meta-analysis-based investigation of acoustic markers in American and Danish autistic children. *Autism Research*, *15*(4), 653–664. https://doi.org/10.1002/aur.2661

Galaz, Z., Mekyska, J., Mzourek, Z., Smekal, Z., Rektorova, I., Eliasova, I., Kostalova, M., Mrackova, M., & Berankova, D. (2016). Prosodic analysis of neutral, stress-modified and rhymed speech in patients with Parkinson's disease. *Computer Methods and Programs in Biomedicine*, *127*, 301-317. https://doi.org/10.1016/j.cmpb.2015.12.011

Goberman, A. M., & Coelho, C. (2002). Acoustic analysis of Parkinsonian speech I: Speech characteristics and L-Dopa therapy. *NeuroRehabilitation*, *17*(3), 237–246. https://doi.org/10.3233/NRE-2002-17310

Goberman, A., Coelho, C., & Robb, M. (2002). Phonatory characteristics of parkinsonian speech before and after morning medication : The ON and OFF states. *Journal of Communication Disorders*, *35*(3), 217-239. https://doi.org/10.1016/s0021-9924(01)00072-7

Goberman, A. M., Coelho, C. A., & Robb, M. P. (2005). Prosodic Characteristics of Parkinsonian Speech : The Effect of Levodopa-Based Medication. *Journal of Medical Speech-Language Pathology*, *13*(1), 51-68.

Goberman, A. M., & Elmer, L. W. (2005). Acoustic analysis of clear versus conversational speech in individuals with Parkinson disease. *Journal of Communication Disorders*, *38*(3), 215-230. https://doi.org/10.1016/j.jcomdis.2004.10.001

Grandjean, D. (2020). Brain Networks of Emotional Prosody Processing. *Emotion Review*, 1754073919898522. https://doi.org/10.1177/1754073919898522

Guenther, F. H. (2016). *Neural Control of Speech*. MIT Press.

Guenther, F. H., & Hickok, G. (2016). Neural Models of Motor Speech Control. In G. Hickok & S. L. Small (Eds.), *Neurobiology of Language* (pp. 725–740). Academic Press. https://doi.org/10.1016/B978-0-12-407794-2.00058-4

Haddaway, N. R., Page, M. J., Pritchard, C. C., & McGuinness, L. A. (2022). PRISMA2020: An R package and Shiny app for producing PRISMA 2020-compliant flow diagrams, with interactivity for optimised digital transparency and Open Synthesis. *Campbell Systematic Reviews*, *18*(2), e1230. https://doi.org/10.1002/cl2.1230

Hansen, L., Rocca, R., Simonsen, A., Parola, A., Bliksted, V., Ladegaard, N., Bang, D., Tylén, K., Weed, E., Østergaard, S. D., & Fusaroli, R. (2023). *Automated speech- and text-based classification of neuropsychiatric conditions in a multidiagnostic setting* (arXiv:2301.06916). arXiv. https://doi.org/10.48550/arXiv.2301.06916

Harel, B. T., Cannizzaro, M. S., Cohen, H., Reilly, N., & Snyder, P. J. (2004). Acoustic characteristics of Parkinsonian speech : A potential biomarker of early disease progression and treatment. *Journal of Neurolinguistics*, *17*(6), 439-453. https://doi.org/10.1016/j.jneuroling.2004.06.001

Hedges, L. V., & Olkin, I. (2014). *Statistical Methods for Meta-Analysis*. Academic Press.

Hickok, G. (2014). Towards an integrated psycholinguistic, neurolinguistic, sensorimotor framework for speech production. *Language, Cognition and Neuroscience*, *29*(1), 52–59. https://doi.org/10.1080/01690965.2013.852907

Hlavnicka, J., Cmejla, R., Tykalova, T., Sonka, K., Ruzicka, E., & Rusz, J. (2017). Automated analysis of connected speech reveals early biomarkers of Parkinson's disease in patients with rapid eye movement sleep behaviour disorder. *Scientific Reports*, *7*, 12. https://doi.org/10.1038/s41598-017-00047-5

Huh, Y. E., Park, J., Suh, M. K., Lee, S. E., Kim, J., Jeong, Y., Kim, H.-T., & Cho, J. W. (2015). Differences in early speech patterns between Parkinson variant of multiple system atrophy and Parkinson's disease. *Brain and Language*, *147*, 14-20. https://doi.org/10.1016/j.bandl.2015.04.007

Illner, V., Tykalová, T., Novotný, M., Klempíř, J., Dušek, P., & Rusz, J. (2022). Toward automated articulation rate analysis via connected speech in Dysarthrias. *Journal of Speech, Language, and Hearing Research*, *65*(4), 1386-1401.

Jacquemot, C., & Bachoud-Lévi, A.-C. (2021). Striatum and language processing: Where do we stand? *Cognition*, *213*, 104785. https://doi.org/10.1016/j.cognition.2021.104785

Jeancolas, L., Mangone, G., Petrovska-Delacrétaz, D., Benali, H., Benkelfat, B.-E., Arnulf, I., Corvol, J.-C., Vidailhet, M., & Lehéricy, S. (2022). Voice characteristics from isolated rapid eye movement sleep behavior disorder to early Parkinson's disease. *Parkinsonism & Related Disorders*, *95*, 86-91. https://doi.org/10.1016/j.parkreldis.2022.01.003

Jeancolas, L., Petrovska-Delacrétaz, D., Mangone, G., Benkelfat, B.-E., Corvol, J.-C., Vidailhet, M., Lehéricy, S., & Benali, H. (2021). X-Vectors: New Quantitative Biomarkers for Early Parkinson's Disease Detection From Speech. *Frontiers in Neuroinformatics*, *15*. https://www.frontiersin.org/articles/10.3389/fninf.2021.578369

Konstantopoulos, K., Vogazianos, P., Christou, Y., & Pisinou, M. (2021). Sequential motion rate and oral reading rate : Normative data for Greek and clinical implications. *Logopedics, Phoniatrics, Vocology*, 1-6. https://doi.org/10.1080/14015439.2021.1901309

Knowles, T., Adams, S. G., & Jog, M. (2021). Variation in Speech Intelligibility Ratings as a Function of Speech Rate Modification in Parkinson's Disease. *Journal of Speech, Language, and Hearing Research*, *64*(6), 1773–1793. https://doi.org/10.1044/2021_JSLHR-20-00593

Kotz, S. A., & Schwartze, M. (2016). Chapter 57 - Motor-Timing and Sequencing in Speech Production: A General-Purpose Framework. In G. Hickok & S. L. Small (Eds.), *Neurobiology of Language* (pp. 717–724). Academic Press. https://doi.org/10.1016/B978-0-12-407794-2.00057-2

Kouba, T., Illner, V., & Rusz, J. (2022). Study protocol for using a smartphone application to investigate speech biomarkers of Parkinson's disease and other synucleinopathies: SMARTSPEECH. *BMJ Open*, *12*(6), e059871. https://doi.org/10.1136/bmjopen-2021-059871

Laganaro, M., Fougeron, C., Pernon, M., Levêque, N., Borel, S., Fournet, M., Chiuvé, S. C., Lopez, U., Trouville, R., Ménard, L., Burkhard, P. R., Assal, F., & Delvaux, V. (2021). Sensitivity and specificity of an acoustic- and perceptual-based tool for assessing motor speech disorders in French: The MonPaGe-screening protocol. *Clinical Linguistics & Phonetics*, *0*(0), 1–16. https://doi.org/10.1080/02699206.2020.1865460

Lam, J., & Tjaden, K. (2016). Clear Speech Variants : An Acoustic Study in Parkinson's Disease. *Journal of Speech, Language, and Hearing Research: JSLHR*, *59*(4), 631-646. https://doi.org/10.1044/2015_JSLHR-S-15-0216

Lee, J., Huber, J., Jenkins, J., & Fredrick, J. (2019). Language planning and pauses in story retell: Evidence from aging and Parkinson's disease. *Journal of Communication Disorders*, *79*, 1–10. https://doi.org/10.1016/j.jcomdis.2019.02.004

Levinson, S. C. (2019). Natural Forms of Purposeful Interaction among Humans What Makes Interaction Effective? In K. A. Gluck & J. E. Laird (Eds.), *Interactive Task Learning: Humans,*

Robots, and Agents Acquiring New Tasks through Natural Interactions. The MIT Press. https://doi.org/10.7551/mitpress/11956.001.0001

Lirani-Silva, C., Mourão, L. F., & Gobbi, L. T. B. (2015). Dysarthria and Quality of Life in neurologically healthy elderly and patients with Parkinson's disease. *CoDAS*, *27*(3), 248-254. https://doi.org/10.1590/2317-1782/20152014083

Lowit, A., & Kuschmann, A. (2012). Characterizing Intonation Deficit in Motor Speech Disorders: An Autosegmental–Metrical Analysis of Spontaneous Speech in Hypokinetic Dysarthria, Ataxic Dysarthria, and Foreign Accent Syndrome. *Journal of Speech, Language, and Hearing Research*, *55*(5), 1472–1484. https://doi.org/10.1044/1092-4388(2012/11-0263)

Lowit, A., Marchetti, A., Corson, S., & Kuschmann, A. (2018). Rhythmic performance in hypokinetic dysarthria: Relationship between reading, spontaneous speech and diadochokinetic tasks. *Journal of Communication Disorders*, *72*, 26–39. https://doi.org/10.1016/j.jcomdis.2018.02.005

Mack, J. E., Chandler, S. D., Meltzer-Asscher, A., Rogalski, E., Weintraub, S., Mesulam, M.-M., & Thompson, C. K. (2015). What do pauses in narrative production reveal about the nature of word retrieval deficits in PPA? *Neuropsychologia*, *77*, 211–222. https://doi.org/10.1016/j.neuropsychologia.2015.08.019

Martens, H., Van Nuffelen, G., Dekens, T., Hernández-Díaz Huici, M., Kairuz Hernández-Díaz, H. A., De Letter, M., & De Bodt, M. (2015). The effect of intensive speech rate and intonation therapy on intelligibility in Parkinson's disease. *Journal of Communication Disorders*, *58*(Supplement C), 91–105. https://doi.org/10.1016/j.jcomdis.2015.10.004

McGrath, S., Zhao, X., Steele, R., Thombs, B. D., Benedetti, A., Levis, B., Riehm, K. E., Saadat, N., Levis, A. W., Azar, M., Rice, D. B., Sun, Y., Krishnan, A., He, C., Wu, Y., Bhandari, P. M., Neupane, D., Imran, M., Boruff, J., … Zhang, Y. (2020). Estimating the sample mean and standard deviation from commonly reported quantiles in meta-analysis. *Statistical Methods in Medical Research*, *29*(9), 2520–2537. https://doi.org/10.1177/0962280219889080

McRae, P. A., Tjaden, K., & Schoonings, B. (2002). Acoustic and perceptual consequences of articulatory rate change in Parkinson disease. *Journal of Speech, Language, and Hearing Research*, *45*(1), 35-50. https://doi.org/10.1044/1092-4388(2002/003)

Meekings, S., & Scott, S. K. (2021). Error in the Superior Temporal Gyrus? A Systematic Review and Activation Likelihood Estimation Meta-Analysis of Speech Production Studies. *Journal of Cognitive Neuroscience*, *33*(3), 422–444. https://doi.org/10.1162/jocn_a_01661

Miller, N. (2017). Communication changes in Parkinson's disease. *Practical Neurology*, practneurol-2017-001635. https://doi.org/10.1136/practneurol-2017-001635

Moreau, C., & Pinto, S. (2019). Misconceptions about speech impairment in Parkinson's disease. *Movement Disorders*, *34*(10), 1471–1475. https://doi.org/10.1002/mds.27791

Parola, A., Simonsen, A., Lin, J. M., Zhou, Y., Wang, H., Ubukata, S., Koelkebeck, K., Bliksted, V., & Fusaroli, R. (2023). Voice Patterns as Markers of Schizophrenia: Building a Cumulative Generalizable Approach Via a Cross-Linguistic and Meta-analysis Based Investigation. *Schizophrenia Bulletin*, *49*(Supplement_2), S125–S141. https://doi.org/10.1093/schbul/sbac128

Pell, M. D., Cheang, H. S., & Leonard, C. L. (2006). The impact of Parkinson's disease on vocal-prosodic communication from the perspective of listeners. *Brain and Language*, *97*(2), 123–134. https://doi.org/10.1016/j.bandl.2005.08.010

Pickering, M. J., & Gambi, C. (2018). Predicting while comprehending language: A theory and review. *Psychological Bulletin*, *144*(10), 1002–1044. https://doi.org/10.1037/bul0000158

Pinto, S., Chan, A., Guimarães, I., Rothe-Neves, R., & Sadat, J. (2017). A cross-linguistic perspective to the study of dysarthria in Parkinson's disease. *Journal of Phonetics*, *64*, 156–167. https://doi.org/10.1016/j.wocn.2017.01.009

Ramanarayanan, V., Lammert, A. C., Rowe, H. P., Quatieri, T. F., & Green, J. R. (2022). Speech as a biomarker: Opportunities, interpretability, and challenges. *Perspectives of the ASHA Special Interest Groups*, *7*(1), 276-283

Ramig, L., Halpern, A., Spielman, J., Fox, C., & Freeman, K. (2018). Speech treatment in Parkinson's disease: Randomized controlled trial (RCT). *Movement Disorders*, *33*(11), 1777–1791. https://doi.org/10.1002/mds.27460

Rektorova, I., Barrett, J., Mikl, M., Rektor, I., & Paus, T. (2007). Functional abnormalities in the primary orofacial sensorimotor cortex during speech in Parkinson's disease. *Movement Disorders*, *22*(14), 2043-2051. https://doi.org/10.1002/mds.21548

Rohatgi, A. (2022). *WebPlotDigitizer User Manual Version 4.6* [Computer software]. https://automeris.io/WebPlotDigitizer

Rusz, J., Cmejla, R., Ruzickova, H., & Ruzicka, E. (2011). Quantitative acoustic measurements for characterization of speech and voice disorders in early untreated Parkinson's disease. *The Journal of the Acoustical Society of America*, *129*(1), 350-367. https://doi.org/10.1121/1.3514381

Rusz, J., Hlavnička, J., Novotný, M., Tykalová, T., Pelletier, A., Montplaisir, J., Gagnon, J.-F., Dušek, P., Galbiati, A., Marelli, S., Timm, P. C., Teigen, L. N., Janzen, A., Habibi, M., Stefani, A., Holzknecht, E., Seppi, K., Evangelista, E., Rassu, A. L., … Šonka, K. (2021). Speech Biomarkers in Rapid Eye Movement Sleep Behavior Disorder and Parkinson Disease. *Annals of Neurology*, *90*(1), 62–75. https://doi.org/10.1002/ana.26085

Rusz, J., Tykalová, T., Krupička, R., Zárubová, K., Novotný, M., Jech, R., Szabó, Z., & Růžička, E. (2017). Comparative analysis of speech impairment and upper limb motor dysfunction in Parkinson's disease. *Journal of Neural Transmission*, *124*(4), 463-470. https://doi.org/10.1007/s00702-016-1662-y

Rusz, J., Tykalová, T., Novotný, M., Růžička, E., & Dušek, P. (2021). Distinct patterns of speech disorder in early-onset and late-onset de-novo Parkinson's disease. *NPJ Parkinson's Disease*, *7*(1), 98. https://doi.org/10.1038/s41531-021-00243-1

Rusz, J., Tykalová, T., Novotný, M., Zogala, D., Růžička, E., & Dušek, P. (2022). Automated speech analysis in early untreated Parkinson's disease : Relation to gender and dopaminergic transporter imaging. *European Journal of Neurology*, *29*(1), 81-90. https://doi.org/10.1111/ene.15099

Rusz, J., Tykalova, T., Ramig, L. O., & Tripoliti, E. (2021). Guidelines for Speech Recording and Acoustic Analyses in Dysarthrias of Movement Disorders. *Movement Disorders*, *36*(4), 803–814. https://doi.org/10.1002/mds.28465

Rybner, A., Jessen, E. T., Mortensen, M. D., Larsen, S. N., Grossman, R., Bilenberg, N., Cantio, C., Jepsen, J. R. M., Weed, E., Simonsen, A., & Fusaroli, R. (2022). Vocal markers of autism: Assessing the generalizability of machine learning models. *Autism Research*, *15*(6), 1018–1030. https://doi.org/10.1002/aur.2721

Sharma, V. K., Singh, T. G., Mehta, V., & Mannan, A. (2023). Biomarkers: Role and Scope in Neurological Disorders. *Neurochemical Research*, *48*(7), 2029–2058. https://doi.org/10.1007/s11064-023-03873-4

Skodda, S., Visser, W., & Schlegel, U. (2010). Short- and long-term dopaminergic effects on dysarthria in early Parkinson's disease. *Journal of Neural Transmission*, *117*(2), 197-205. https://doi.org/10.1007/s00702-009-0351-5

Skodda, S., Visser, W., & Schlegel, U. (2011). Gender-Related Patterns of Dysprosody in Parkinson Disease and Correlation Between Speech Variables and Motor Symptoms. *Journal of Voice*, *25*(1), 76-82. https://doi.org/10.1016/j.jvoice.2009.07.005

Skrabal, D., Tykalova, T., Klempir, J., Ruzicka, E., & Rusz, J. (2020). Dysarthria enhancement mechanism under external clear speech instruction in Parkinson's disease, progressive supranuclear palsy and multiple system atrophy. *Journal of Neural Transmission*, *127*(6), 905-914. https://doi.org/10.1007/s00702-020-02171-5

Tanaka, Y., Nishio, M., & Niimi, S. (2011). Vocal acoustic characteristics of patients with Parkinson's disease. *Folia Phoniatrica et Logopaedica: Official Organ of the International Association of Logopedics and Phoniatrics (IALP)*, *63*(5), 223-230. https://doi.org/10.1159/000322059

Thies, T., Mücke, D., Lowit, A., Kalbe, E., Steffen, J., & Barbe, M. T. (2020). Prominence marking in parkinsonian speech and its correlation with motor performance and cognitive abilities. *Neuropsychologia*, *137*, 107306. https://doi.org/10.1016/j.neuropsychologia.2019.107306

Tjaden, K., & Wilding, G. (2011a). Speech and pause characteristics associated with voluntary rate reduction in Parkinson's disease and Multiple Sclerosis. *Journal of Communication Disorders*, *44*(6), 655-665. https://doi.org/10.1016/j.jcomdis.2011.06.003

Tjaden, K., & Wilding, G. (2011b). The impact of rate reduction and increased loudness on fundamental frequency characteristics in dysarthria. *Folia Phoniatrica et Logopaedica: Official Organ of the International Association of Logopedics and Phoniatrics (IALP)*, *63*(4), 178-186. https://doi.org/10.1159/000316315

Tjaden, K., & Wilding, G. E. (2004). Rate and loudness manipulations in dysarthria : Acoustic and perceptual findings. *Journal of Speech, Language, and Hearing Research: JSLHR*, *47*(4), 766-783. https://doi.org/10.1044/1092-4388(2004/058)

Tjaden, K., & Wilding, G. E. (2005). Effect of Rate Reduction and Increased Loudness on Acoustic Measures of Anticipatory Coarticulation in Multiple Sclerosis and Parkinson's Disease. *Journal of Speech, Language, and Hearing Research*, *48*(2), 261-277. https://doi.org/10.1044/1092-4388(2005/018)

Tykalova, T., Rusz, J., Cmejla, R., Ruzickova, H., & Ruzicka, E. (2014). Acoustic Investigation of Stress Patterns in Parkinson's Disease. *Journal of Voice*, *28*(1), 129.e1-129.e8. https://doi.org/10.1016/j.jvoice.2013.07.001

Tykalová, T., Rusz, J., Švihlík, J., Bancone, S., Spezia, A., & Pellecchia, M. T. (2020). Speech disorder and vocal tremor in postural instability/gait difficulty and tremor dominant subtypes of Parkinson's disease. *Journal of Neural Transmission*, *127*(9), 1295-1304. https://doi.org/10.1007/s00702-020-02229-4

Thornton, A., & Lee, P. (2000). Publication bias in meta-analysis: Its causes and consequences. *Journal of Clinical Epidemiology*, *53*(2), 207–216. https://doi.org/10.1016/S0895-4356(99)00161-4

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

Van Lancker Sidtis, D., & Yang, S. (2020). Pathological Prosody: Overview, assessment, and treatment. In C. Gussenhhoven & A. Chen (Eds.), *The Oxford Handbook of Language Prosody* (p. 0). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780198832232.013.48

Vaysse, R., Astésano, C., & Farinas, J. (2022). Performance analysis of various fundamental frequency estimation algorithms in the context of pathological speech. *The Journal of the Acoustical Society of America*, *152*(5), 3091–3101. https://doi.org/10.1121/10.0015143

Vaysse, R., Ghio, A., Astésano, C., Farinas, J., & Viallet, F. (2022). Analyse macroscopique des variations et modulations de F0 en lecture dans la maladie de Parkinson : Données sur 320 locuteurs. *XXXIVe Journées d'Études sur la Parole -- JEP 2022*, 307–315. https://doi.org/10.21437/JEP.2022-33

Verkhodanova, V., Coler, M., Jonkers, R., Timmermans, S., Maurits, N., de Jong, B., & Lowie, W. (2022). A cross-linguistic perspective to classification of healthiness of speech in

Parkinson's disease. *Journal of Neurolinguistics*, *63*, 1-20. https://doi.org/10.1016/j.jneuroling.2022.101068

Viechtbauer, W. (2010). Conducting Meta-Analyses in R with the metafor Package. *Journal of Statistical Software*, *36*, 1–48. https://doi.org/10.18637/jss.v036.i03

Viñas-Guasch, N., & Wu, Y. J. (2017). The role of the putamen in language: A meta-analytic connectivity modeling study. *Brain Structure and Function*, *222*(9), 3991–4004. https://doi.org/10.1007/s00429-017-1450-y

Wagenmakers, E.-J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., Love, J., Selker, R., Gronau, Q. F., Šmíra, M., Epskamp, S., Matzke, D., Rouder, J. N., & Morey, R. D. (2018). Bayesian inference for psychology. Part I: Theoretical advantages and practical ramifications. *Psychonomic Bulletin & Review*, *25*(1), 35–57. https://doi.org/10.3758/s13423-017-1343-3

Walsh, B., & Smith, A. (2012). Basic parameters of articulatory movements and acoustics in individuals with Parkinson's disease. *Movement Disorders*, *27*(7), 843-850. https://doi.org/10.1002/mds.24888

Whitehill, T. L. (2010). Studies of Chinese speakers with dysarthria: informing theoretical models. *Folia Phoniatrica et Logopaedica*, *62*(3), 92-96.

Whitfield, J. A., & Gravelin, A. C. (2019). Characterizing the distribution of silent intervals in the connected speech of individuals with Parkinson disease. *Journal of Communication Disorders*, *78*, 18-32. https://doi.org/10.1016/j.jcomdis.2018.12.001

Whitehill, T. L., Kwan, L., Lee, F. P.-H., & Chow, M. M.-N. (2011). Effect of LSVT on Lexical Tone in Speakers with Parkinson's Disease. *Parkinson's Disease*, *2011*, e897494. https://doi.org/10.4061/2011/897494

Chapitre 2
Effets de l'amorçage rythmique sur la production de la parole dans la maladie de Parkinson et l'ataxie spinocérébelleuse

Le présent chapitre prolonge les analyses présentées dans le chapitre précédent en examinant si une amorce rythmique est susceptible de modifier la production de la parole.

Ce travail est présenté dans cette thèse sous la forme d'un article intitulé *When rhythm guides speech production: insights from Parkinson's disease and spinocerebellar ataxia*, actuellement en révision. L'étude repose sur un paradigme d'amorçage rythmique dans lequel les participants devaient lire à voix haute des phrases immédiatement après avoir entendu une brève amorce rythmique non verbale, ou après une période de silence de durée équivalente. Les amorces pouvaient être régulières, c'est-à-dire organisées selon une structure rythmique correspondant à celle des phrases cibles, ou irrégulières, sans organisation métrique stable.

L'étude compare trois groupes de participants : des participants contrôles, des personnes atteintes de la MP et des personnes atteintes d'ASC. Ce choix permet d'examiner les effets de l'amorçage rythmique dans deux populations présentant des troubles de la parole, mais aussi des atteintes neurofonctionnelles distinctes, impliquant notamment les réseaux cortico-sous-corticaux et cérébelleux engagés dans le traitement temporel et la production motrice de la parole. L'objectif n'est donc pas seulement de déterminer si le rythme facilite la production, mais aussi d'examiner si cette facilitation varie selon le profil clinique des participants.

Plusieurs dimensions de la production de la parole ont été analysées : l'exactitude de la production, le temps d'initiation de la parole, le débit de parole, ainsi qu'un ensemble de métriques rythmiques. Ces métriques reposent sur l'analyse des durées vocaliques,

consonantiques et syllabiques, et permettent de quantifier différents aspects de l'organisation temporelle de la parole, tels que la proportion de temps vocalique, la variabilité des intervalles ou les différences de durée entre intervalles successifs.

Les résultats montrent que l'effet le plus robuste concerne le temps d'initiation : les amorces régulières raccourcissent de manière crédible le délai d'entrée en parole dans les trois groupes. En revanche, les effets observés sur le débit, l'exactitude et les métriques rythmiques apparaissent plus limités. Ce profil de résultats suggère que, dans ce paradigme, le bénéfice principal d'un rythme externe porte avant tout sur la préparation temporelle globale de la réponse parlée, plutôt que sur une modification systématique de l'organisation temporelle fine de l'énoncé.

Le résumé en français et l'article dans sa version en révision sont présentés ci-après.

Fumel, J., Mairano, P., Ott, L., Contreras Roa, L., Devos, D., Moreau, C., & Basirat, A. (in revision). *When rhythm guides speech production: insights from Parkinson's disease and spinocerebellar ataxia*.

Résumé

L'exposition auditive rythmique module le traitement de la parole qui suit, mais son effet sur la production de la parole reste mal compris. Les personnes atteintes de la maladie de Parkinson (MP) ou d'ataxie spinocérébelleuse (ASC) souffrent de dysarthrie, pour laquelle il n'existe aucun traitement. À l'aide d'une approche d'amorçage rythmique, nous avons cherché à déterminer si de brefs stimuli rythmiques modulent la production de la parole et la facilitent chez les personnes atteintes de MP ou d'ASC, considérées ici comme des modèles cliniques de dysfonctionnement, respectivement, des ganglions de la base et du cervelet. Cinquante-huit locuteurs natifs du français (20 témoins, 24 personnes atteintes de MP, 14 personnes atteintes d'ASC) ont lu des phrases après un stimulus rythmique régulier, un stimulus irrégulier ou un silence. Nous avons mesuré la précision et les caractéristiques temporelles de la production de la parole, notamment le temps d'initiation de la parole, le débit de parole et les métriques rythmiques. Des modèles bayésiens à effets mixtes ont été utilisés pour l'analyse statistique de ces résultats. La précision n'a montré aucun effet de l'amorçage rythmique. Les amorces régulières ont systématiquement raccourci le temps d'initiation de la parole dans toutes les populations, mais le bénéfice des amorces régulières était plus faible chez les personnes atteintes de la MP que chez les témoins. Chez les personnes atteintes d'ASC, le débit de parole a diminué après les amorces régulières. L'impact de l'amorçage rythmique sur les métriques rythmiques s'est avéré limité. Dans l'ensemble, ces résultats indiquent que l'amorçage rythmique influence principalement le contrôle global plutôt que fin du timing de la parole lors de la production vocale. Ils mettent également en lumière le rôle des ganglions de la base et du cervelet dans la modulation de la parole par le rythme. Ils apportent en outre des éclaircissements sur les modèles de production de la parole, qui ne disposent actuellement d'aucun mécanisme explicite pour traiter les entrées rythmiques externes. Sur le plan clinique, ils contribuent à éclairer les approches orthophoniques basées sur le rythme.

Mots-clés : production de la parole, rythme, lien perception-production, ganglions de la base, cervelet

When rhythm guides speech production: insights from Parkinson's disease and spinocerebellar ataxia

Jules Fumel[1*], Paolo Mairano[2], Laurent Ott[1], Leonardo Contreras Roa[3], David Devos[4], Caroline Moreau[4] & Anahita Basirat[1]

[1]Univ. Lille, CNRS, UMR 9193—SCALab—Sciences Cognitives et Sciences Affectives, Lille, France

[2]Univ. Tours, CNRS, UMR 7270 – LLL –Laboratoire Ligérien de Linguistique, F-37000 Tours, France

[3]EA 4295 Corpus, Université de Picardie Jules Verne, 80080 Amiens, France

[4]Univ. Lille, Inserm, CHU Lille, U1172 TREAT Team—LilNCog—Lille Neuroscience & Cognition, Lille, France

Abstract

Rhythmic auditory exposure modulates subsequent speech processing, yet its effect on speech production remains poorly understood. People with Parkinson's disease (PD) or spinocerebellar ataxia (SCA) suffer from dysarthria, for which there is no treatment. Using a rhythmic priming approach, we investigated whether brief rhythmic primes modulate speech production and facilitate it in people with PD and SCA considered here as clinical models of basal ganglia dysfunction and cerebellar dysfunction, respectively. Fifty-eight native French speakers (20 controls, 24 people with PD, 14 people with SCA) read sentences following a regular prime, an irregular prime, or silence. We measured accuracy and the temporal features of speech production including speech initiation time, speech rate, and rhythmic metrics. Bayesian mixed-effects models were used for the statistical analyses of these outputs. Accuracy showed no effect of rhythmic priming. Regular primes consistently shortened speech initiation time in all populations, but the benefit of regular primes was smaller in people with PD than in controls. In People with SCA, speech rate decreased after regular primes. The impact of rhythmic priming on rhythmic metrics was limited. Together, these results indicate that perception-to-production rhythmic priming mainly affects global, rather than fine-grained, timing control of speech during speech production. They also shed light on the implications of the basal ganglia and cerebellum in the modulation of speech by rhythm. They also provide insights for speech production models, which currently lack an explicit mechanism for external rhythmic inputs. Clinically, they help inform rhythm-based speech and language therapy methods.

Keywords: Speech production, rhythm, perception-production link, basal ganglia, cerebellum

1 Introduction

A growing body of evidence suggests that exposure to rhythmic auditory patterns can modulate the processing of subsequent speech. This effect, known as rhythmic priming, has been observed at various levels of processing such as phonological and syntactic aspects, across tasks and populations (Bedoin et al., 2016; Canette et al., 2019; Canette et al., 2020a; Cason & Schön, 2012; Cason et al., 2015a; Chern et al., 2018; Falk et al., 2017; Fiveash et al., 2020; György et al., 2024; Kotz & Gunter, 2015; Ladányi et al., 2021; Przybylski et al., 2013). In rhythmic priming paradigms, participants are first exposed to auditory rhythmic patterns, after which their performance on specific speech and language processing tasks is assessed. To examine the impact of rhythm, performance following regular rhythmic primes is typically contrasted with that observed in control conditions, such as no-priming baselines or primes consisting of irregular rhythmic patterns. The findings point to a beneficial effect of regular primes over control conditions, suggesting that the beneficial effect is due to better performance with regular primes and not to potential processing costs associated with irregular primes (Fiveash et al., 2025). The majority of rhythmic priming studies have focused on speech perception and comprehension, whereas the impact of rhythmic priming on speech production has been investigated less extensively. Questions remain regarding the extent to which perception-production transfer may occur, as well as the neurocognitive mechanisms underlying this effect. The present study aimed to address this gap by examining the impact of rhythmic priming on various aspects of speech production. It also assessed the relevance of rhythmic priming in Parkinson's disease (PD) and spinocerebellar ataxia (SCA), both of which are associated with speech impairments with severe consequences on autonomy (Atkinson-Clement et al., 2019; Walshe & Miller, 2011). Clarifying how rhythmic priming influences speech production in these disorders could shed light on the underlying mechanisms of this effect and guide potential clinical applications in speech and language pathology.

1.1 Impact of rhythmic priming on speech and language processing

The rhythmic priming effect is typically explained within the framework of Dynamic Attending Theory (DAT) (Large & Jones, 1999; Jones, 2019). According to DAT, when an external rhythm is presented, the internal oscillations, known as attending rhythms, entrain to that external rhythm. This is believed to optimize perception by directing attention to

relevant points in time, enhancing the salience of an event's temporal structure, facilitating predictions about upcoming events, and thereby improving performance. This effect continues even after the external stimulus has ceased, a phenomenon known as the echo effect in the domain of speech processing (Kösem et al., 2018; Zoefel et al., 2024). The Processing Rhythm in Speech and Music (PRISM) framework is also a relevant model for understanding the rhythmic priming effect. PRISM suggests three common mechanisms underlying rhythm processing in both music and speech (Fiveash et al., 2021): (i) Sensitivity to detailed variations in rhythm, pitch, or timbre; (ii) alignment of neural oscillations with external rhythmic patterns, a process referred to as neural entrainment; and (iii) sensorimotor coupling, engaging the interaction between the auditory and motor cortices. In line with these frameworks, rhythmic priming is thought to operate through neural entrainment to rhythmic patterns, which enhances the formation of temporal expectations about upcoming events and thereby facilitates their processing. In addition, although rhythmic priming does not necessarily require the involvement of the motor system, its activation through sensorimotor synchronization, such as tapping along with the rhythmic prime, has been shown to enhance the effect of rhythmic priming on speech processing (Cason et al., 2015a). The underlying hypothesis is that recruiting the motor system in such tasks can improve the temporal precision of auditory attention (Berthault et al., 2024). In the present study, we propose that the principles of the DAT and PRISM frameworks also extend to speech production, such that rhythmic priming may enhance speech motor timing.

1.2 Rhythm perception-production transfer and speech production

A few studies to date have suggested that prior exposure to regular rhythms can influence speech production. Rhythmic priming has been reported to improve phonological production accuracy in hearing-impaired children (Cason et al., 2015b) and sentence repetition, indexed by grammatical accuracy, in children with developmental language disorders (Fiveash et al., 2023). Importantly, with regard to the aim of the current study, rhythmic priming has also been shown to influence overall speech timing in adults and children, with faster musical primes leading to faster subsequent speech production (Jungers & Hupp, 2018). Exposure to rhythmic primes has also been reported to reduce speech initiation time in adults (Zhang & Zhang, 2019), and this facilitation may be further enhanced

when listening to the prime is combined with active motor engagement (Gindre, 2024). Finally, in adult second-language learners, rhythmic primes have been shown to make learners' speech production more similar to the rhythm of the target language (Liu & Liu, 2024). These findings support the plausibility of cross-domain transfer from nonverbal rhythm to the linguistic processes involved in speech and language production. They can be interpreted within the framework of DAT: a rhythmic auditory stimulus may engage domain-general timing and prediction processes and recruit an auditory–motor network, which in turn influences the timing of speech production, affecting both the initiation of speech and its temporal unfolding.

Regarding the neural substrates, insights into these rhythmic perception–production priming effects can be drawn from literature examining the role of the basal ganglia and cerebellum in temporal processing (Teki et al., 2011; Teki et al., 2012), including their contributions to speech timing (Kotz & Schwartze, 2010; Kotz & Schwartze, 2016). The basal ganglia are thought to be crucial for the internal generation of a periodic beat and the prediction of future events through a relative timing mechanism (Grahn, 2009; Grahn & Rowe, 2009). In contrast, the cerebellum is believed to specialize in absolute duration-based mechanisms, encoding the precise timing of individual events (Grube et al., 2010; Breska & Ivry, 2018; Breska & Ivry, 2021). In a study on the neural tracking of rhythm, Nozaradan et al. (2017) compared individuals with focal lesions in the basal ganglia or in the cerebellum, and healthy controls. The individuals with lesions displayed neural responses to rhythmic sequences; however, these responses were reduced compared to those of the control group. Importantly, the pattern of results diverged across the two clinical groups: individuals with basal ganglia lesions exhibited atypical neural responses to complex rhythmic structures, whereas individuals with cerebellar lesions showed diminished neural tracking at faster temporal rates. Although these studies did not directly examine the rhythmic priming effect, they nonetheless suggest that both the basal ganglia and the cerebellum may contribute to rhythmic priming, albeit in different ways.

Only a limited number of studies to date have examined how disruptions in these structures might modulate the rhythmic priming effect in relation to speech and language processing. In their case study, Kotz and Gunter (2015) found that rhythmic priming enhanced subsequent syntactic and semantic processing in people with PD (PwPD). To our knowledge,

the only study examining the role of the basal ganglia and cerebellum in the effect of perceived rhythm on speech production is that of Späth et al. (2022). In this study, PD was considered as a clinical model of basal ganglia dysfunction, and SCA served as a model of cerebellar dysfunction. Participants listened to sentences with regular or irregular rhythms before producing speech. The authors analyzed the extent to which participants unintentionally adjusted their speech rhythm to match that of the presented sentences produced by a model speaker, a phenomenon called adaptation or convergence (Pardo, 2013). The results showed that people with SCA (PwSCA) exhibited a markedly reduced rhythmic adaptation effect, whereas PwPD showed a pronounced effect. In this study, rhythm was conveyed through speech, distinguishing the paradigm from rhythmic priming studies that used nonverbal primes. Nevertheless, these results can still suggest that the basal ganglia and cerebellum may contribute differently to the benefits of perceived rhythm for speech production, with PwPD showing a greater degree of rhythmic adaptation.

Interestingly, findings from studies on rhythmic auditory stimulation suggest that rhythm has beneficial effects in PD. Rhythmic auditory stimulation involves the use of external, non-verbal rhythmic cues while individuals engage in movements such as walking. Strong evidence indicates that PwPD respond positively to these rhythmic interventions (Dalla Bella et al., 2015; Ghai et al., 2018; Nombela et al., 2013; Thaut et al., 2001). One explanation is that auditory rhythm promotes a compensatory mechanism for rhythm processing by enhancing the contribution of the cerebello–thalamo–cortical circuit. Another explanation is that the beneficial effect depends on residual activity within the basal ganglia–thalamo–cortical circuit, whereby activation of auditory regions through rhythmic stimulation might influence this network. Although current evidence in the literature is insufficient to distinguish between these possibilities (Koshimori & Thaut, 2018), it suggests that rhythmic priming may represent a promising therapeutic approach.

1.3 Current study

The current study aimed to examine whether and how rhythmic auditory primes facilitate subsequent sentence production. To this end, we tested the impact of three different prime types on speech production: regular primes reflecting the structure of the sentences to be produced, irregular primes with no specific rhythmic structure, and silence. Three groups

of participants were included: PwSCA, PwPD, and neurologically healthy participants (CTRL). PD and SCA are both known to cause speech impairments. In PD, hypokinetic dysarthria typically manifests as monopitch and monoloudness, reduced stress, imprecise consonants, variable rate with short rushes, and inappropriate silences (Darley et al., 1969; Duffy, 2020; Fumel et al., 2024). In SCA, ataxic dysarthria presents features from the clusters defined by Darley et al. (1969): prosodic excess (excess and equal stress; prolonged phonemes and intervals), articulatory inaccuracy (imprecise consonants, irregular articulatory breakdowns, distorted vowels), and phonatory–prosodic insufficiency (harsh voice, monopitch, slow rate) (see also Duffy, 2020; Schalling & Hartelius, 2013). It is noteworthy that impairments extend beyond these features and also affect other aspects of speech and language processing (for reviews, see Basirat, 2025; Silveri, 2021). The current study aimed to explore whether rhythmic priming could mitigate speech-related symptoms. It is noteworthy that the study of Späth et al. (2022) examined rhythmic adaptation by having PwPD and PwSCA listen to sentences with regular or irregular rhythmic structures produced by a model speaker before speaking, rather than using nonverbal rhythmic stimuli. This left open the question of how purely musical rhythms might affect speech production in PwPD and PwSCA.

We examined several speech outcomes: accuracy, initiation time, speech rate, and rhythmic metrics. Rhythmic metrics were originally developed to distinguish between languages exhibiting different rhythmic structures in a systematic way (e.g., Ramus et al., 1999). Although their use in capturing cross-linguistic differences has been criticized (Arvaniti, 2012), rhythmic metrics have been effectively applied to describe distinctive speech patterns associated with various forms of dysarthria, including hypokinetic and ataxic dysarthrias (e.g., Liss et al., 2009; Lowit, 2014; Lowit et al., 2018). In addition to the primary aims outlined above, the impact of rhythm on speech processing may be associated with musical training and expertise (Patel, 2014). A secondary goal of the study was to examine whether this holds true for the impact of rhythmic priming on speech production.

Based on the evidence reviewed above, we formulated the following hypotheses regarding the effect of regular primes compared to silence and irregular primes in our study. Across all groups, we expected regular primes to accelerate speech initiation (H1). In the clinical groups, we expected regular primes to normalize speech rate (H2) as well as accuracy and rhythmic metrics (H3) in PwPD, but not or only minimally in PwSCA. Finally, across all

groups, we hypothesized that the magnitude of rhythmic priming effects would correlate with musical experience (H4).

2 Methods

2.1 Participants

Twenty-four PwPD, 14 PwSCA and 20 CTRL participated in this study. The study was conducted in accordance with the tenets of the Helsinki Declaration and approved by the local data protection committee. Before testing, all participants were informed about the study. They signed an informed consent form at the beginning of the study. Table 1 shows the characteristics of the participants. All participants were native French speakers and reported normal or corrected-to-normal vision, with no self-reported history of psychiatric or neurological disorders, other than PD in PwPD and SCA for PwSCA. None of the participants exhibited depression symptoms exceeding the mild level, as assessed by the Geriatric Depression Scale (Yesavage et al., 1983), and none were professional musicians. Cognitive performance was assessed using the Montreal Cognitive Assessment (MoCA; Nasreddine et al., 2005), a widely used screening tool for detecting cognitive impairment across various clinical populations, including PwPD and PwSCA (e.g., Dalrymple-Alford et al., 2010; De Picciotto et al., 2024). The cut-off was set to 22. Hearing thresholds were assessed using a screening audiometer. An average hearing threshold below 40 dB HL across the frequencies of 500 Hz, 1000 Hz, and 2000 Hz was considered acceptable. This threshold corresponds to the lower limit of moderate hearing loss. For participants with PD and SCA, inclusion required a neurologist-confirmed diagnosis of idiopathic PD without deep brain stimulation and SCA, respectively. Participants with PD were tested in the ON-medication state. PwSCA presented with a predominantly cerebellar syndrome, characterized by postural and gait disturbances and limb ataxia. Disease severity was assessed using the motor subscore of the Unified Parkinson's Disease Rating Scale (UPDRS; Goetz et al., 2008) for PwPD and the Scale for the Assessment and Rating of Ataxia (SARA; Schmitz-Hübsch et al., 2006) for PwSCA. These scores could not be retrieved for five participants, one in PD and four in SCA populations. Speech impairment was assessed by a speech-language pathologist for all participants using the *Batterie d'Évaluation Clinique de la Dysarthrie* (BECD; Auzou & Rolland-Monnoury, 2019). The perceptual score of BECD was derived from a text-reading task. This score provides an index

of dysarthria severity ranging from 0 (0: normal speech) to 20 (very severe dysarthria). Finally, musical experience was assessed via a self-report questionnaire covering musical training and practice, with scores ranging from 0 to 23, where 23 reflects a high level of past or current musical engagement. This measure is hereafter referred to as the musicality score.

Table 1. Demographic and clinical data of the experimental populations.

	CTRL (n=20)	PD (n=24)	SCA (n=14)	Between-population differences[a]
Gender	11 f, 9 m	12 f, 12 m	11 f, 3 m	$\chi^2(2) = 3.14$
Age (years)	66.50 [48, 81]	67 [50, 80]	63 [48, 76]	H = 3.34
Education duration (years)	14.50 [7, 20]	13 [7, 17]	13.5 [9, 19]	H= 2.68
Hearing threshold (dB)	16.5 [6, 39.50]	17.25 [4.5,30.5]	18.75 [7, 38]	H = 0.20
MOCA	28 [23, 30]	28 [23, 30]	28 [22, 30]	H = 0.18
Dysarthria severity score	1 [0, 7]	6.5 [1, 16]	13 [1, 18]	H = 32.57** SCA > CTRL, PD > CTRL
Musicality Score	4 [0, 20]	2 [0, 14]	0 [0, 9]	H = 5.40
Disease Duration (years)	-	6.5 [1, 18]	18.5 [0, 30]	U = 250* SCA > PD
SARA	-	-	18.5 [4, 33][b]	-
UPDRS-III	-	13 [6, 35][c]	-	-

Notes. Values are median [min-max]. CTRL: Controls; PD: Parkinson's disease; SCA: Spinocerebellar ataxia; MOCA: Montreal Cognitive Assessment; SARA: Scale for the Assessment and Rating of Ataxia; UPDRS-III: Unified Parkinson's Disease Rating Scale – motor subscore.

[a]Between-population differences (χ^2, Kruskal-Wallis-H, Mann-Whitney-U). *p<0.05, **p<0.001. [b]n = 10, data from 4 participants could not be retrieved. [c]n = 23, data from 1 participant could not be retrieved.

2.2 Materials

Forty-five French sentences were created for this study. All sentences were declarative and composed of four Accentual Phrases (APs), each containing three syllables. In French, the AP is the lowest level among prosodic hierarchical levels and can contain more than one lexical word (e.g., [*Européen*] [*est un mot*] [*utilisé*] [*par les français*] (European is a word used by French people, Jun & Fougeron, 2000, p. 215), "[]" indicate the boundaries of the AP). The right boundary of an AP is marked by a perceptually prominent syllable, called AP-final syllable. The temporal organization of each sentence in terms of non-prominent and prominent syllables was thus "xxX xxX xxX xxX" (x: non-prominent, X: prominent). The four APs were grouped into two successive intermediate phrases (ip). For example: {ip [AP *j'ai mangé*] [AP *des croissants*] } {ip [AP *au goûter*] [AP *aujourd'hui]* } (I ate croissants for snack time today). Lexical items such as monosyllabic or disyllabic adjectives, for example, *grand* (big), *méchant* (naughty), as well as adverbs likely to induce semantic emphasis such as *jamais* (never), commonly associated with focus accentuation (Astésano & Bertrand, 2016; Delais, 1993), were excluded, as they could disrupt the intended metrical structure. From a syntactic standpoint, the sentences were kept simple: passive constructions, complex clauses, and modal verbs were deliberately excluded. In addition to the 45 target sentences, nine distractor sentences were selected from the FHarvard corpus (Aubanel et al., 2020). These sentences were all in the infinitive form and followed a different metrical structure to prevent habituation to either rhythmic or syntactic patterns across trials. The 54 total sentences (i.e., 45 targets and nine distractors) were divided into three lists. Each list included 15 target sentences and three distractors. Sentences were paired across the three lists based on several criteria: Word and syllable length, lexical frequency derived from Lexique 3.6 database (New et al., 2001), morphosyntactic characteristics (semantic category of the subject noun, verb tense, gender and number of the subject pronoun).

Rhythmic stimuli were created using the Hydrogen digital audio software (Cominu, 2021), and exported in WAV format with a 44,100 Hz sampling rate. We first generated a regular rhythmic pattern aligned with the metrical structure of the target sentences ("xxX xxX xxX xxX"). Two types of percussive sounds were used to distinguish between unaccented syllables ("x") and accented syllables ("X"), both differing in duration and timbre. We used percussive sounds, rather than tonal instruments, to avoid introducing melodic cues and to

maintain a purely rhythmic structure. The accented beat lasted 150 ms followed by 150 ms of silence, whereas the unaccented beat lasted 100 ms followed by 50 ms of silence. For both beats, we quantified the intensity on the sound portion excluding the subsequent silence. We also calculated the spectral center of gravity, a measure that can contribute to timbre characterization (Peeters et al., 2011), on the 30 ms onset window of the beats. These analyses were performed using Praat (version 6.4.19) (Boersma, 2024). The accented sound was 8 dB more intense on average than the unaccented sound (mean intensity: 73 dB vs. 65 dB) and also showed a higher peak intensity (84 dB vs. 77 dB). The spectral center of gravity was 791.7 Hz and 757.6 Hz for accented and unaccented sounds, respectively. In addition to the regular prime, three irregular rhythmic primes were created by randomly altering the position of stressed and unstressed beats. Thus, the regular prime was congruent with the rhythmic structure of target sentences, and the irregular primes were incongruent. Each auditory prime had a fixed duration of 2.4 seconds.

2.3 Procedure

We used a rhythmic priming paradigm. Each trial consisted of reading a sentence after listening to the regular auditory stimuli (Regular condition), listening to one of the irregular auditory stimuli (Irregular condition), or experiencing a silence period matching the duration of auditory primes (Silence condition). Figure 1 shows the trial timeline. The experiment was divided into three blocks, with trials in each block sharing the same condition type. Each block comprised one sentence list containing 15 target sentences and three distractors. Each participant was presented with all 54 sentences. In the irregular block, each irregular auditory stimulus was presented five times with the target sentences. Within each block, sentence presentation order was randomized. Across participants, each sentence list was paired with all priming conditions. During the experiment, participants were seated approximately 50 cm from the screen and wore a Shure SM35 headset microphone connected to a MOTU M2 audio interface. The program for stimuli presentation and speech recording was developed in GNU Octave (version 5.2.0) using the Psychophysics Toolbox Version 3 (PTB-3 version 3.0.17) and was run on Ubuntu 20.04.3. Each trial started with a white fixation cross on a black background. The auditory stimuli were played free-field through the computer speakers. Immediately after the prime period, a sentence appeared and participants read it aloud while

speech was recorded. They advanced to the next trial with a key-press. If no response occurred, the trial ended automatically after 15 s. A short break was offered every six sentences.

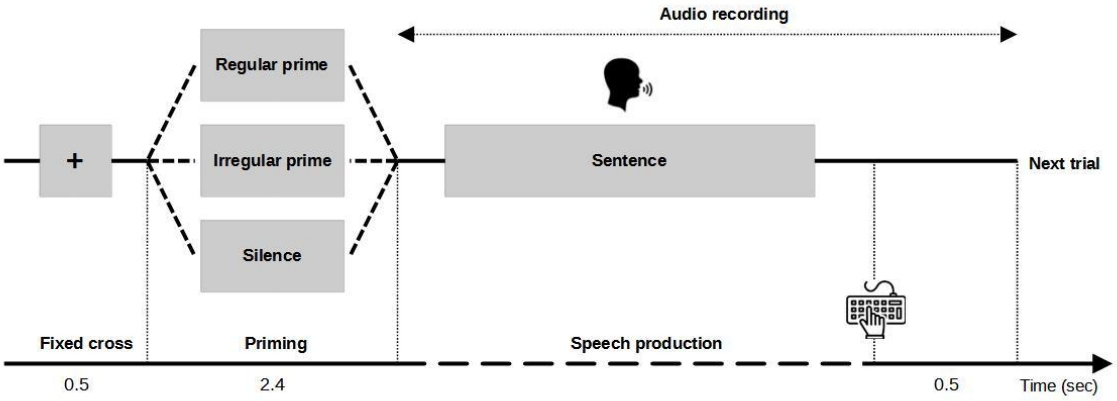

Figure 1. Trial timeline used in this study.

2.4 Data analysis

The audio files were automatically transcribed and segmented at phoneme and syllable levels using WebMAUS (Kisler et al., 2017). After this step, all transcriptions and boundaries were manually reviewed and corrected using Praat software (version 6.4.19) (Boersma, 2024), to address potential errors in transcription accuracy, temporal alignment, and syllable segmentation. Silent intervals of at least 50 ms within utterances were annotated as pauses. Pause annotation was necessary to accurately segment vocalic, consonantal, and syllabic intervals, as their durations were used in our analysis (see below). Pauses themselves were not analyzed as a speech outcome in the present study. For each audio file, we measured the initiation time, defined as the duration between the end of the priming period and the onset of speech production, and the speech rate, obtained by dividing the number of syllables in a sentence by its total duration including pauses. We also computed rhythmic metrics based on the duration of vocalic, consonantal and syllabic intervals. For each sentence, we calculated the proportion of time devoted to vowels, out of the total duration of vocalic and consonantal intervals, expressed in percentage (%V). The ordered sequences of interval durations of vocalic, consonantal and syllabic intervals were then used to calculate the following measures

related to duration variability: ΔV, ΔC, ΔS, VarcoV, VarcoC, VarcoS, nPVI-V, rPVI-C and nPVI-S, where V, C and S stand for vocalic, consonantal and syllabic intervals, respectively, Δ for standard deviation, Varco for coefficient of variation, and PVI measures quantify average absolute differences between successive intervals (normalized for nPVI, raw for rPVI). It is noteworthy that we used both phonemic and syllabic metrics, as they capture complementary temporal scales. This two-level approach is commonly employed in dysarthria research, including studies on PD and SCA (e.g., Lowit, 2014; Lowit et al., 2018; Liss et al., 2009). A summary of the measures used in the study is provided in Table 2.

A total of 2610 target sentences were collected, some of which were excluded to ensure robust speech analyses based on two criteria. First, sentences affected by technical issues that prevented proper recording of the utterance were excluded. Second, sentences containing segmental errors or disruptive events (e.g., coughing) were excluded. With regard to errors, sentences were excluded if they did not match the expected target with the 12 expected syllables. This restriction was applied because segmental errors, including missing or added phonemes and syllables, render speech outcomes non-comparable across experimental conditions (see also Späth et al., 2022, for a similar approach in their study on speech rhythm). To limit the number of exclusions, sentences in which the detected issue occurred only after the first syllable were retained for speech initiation time analyses, but excluded from speech rate and rhythmic metric analyses. Applying the first criterion, one sentence was excluded from all analyses and an additional 32 sentences were excluded from the speech rate and rhythmic metric analyses. Applying the second criterion, 97 sentences were excluded from the speech initiation time analysis and 464 sentences from the speech rate and rhythmic metric analyses. Thus, in total, 2512 sentences were retained for the analysis of speech initiation time (876 for CTRL, 1044 for PwPD, and 592 for PwSCA) and 2113 sentences for the analyses of speech rate and rhythmic metrics (766 for CTRL, 893 for PwPD, and 454 for PwSCA). We then calculated an accuracy score defined as a binary sentence-level outcome: sentences were labeled as correct when the target sentence was produced as intended in terms of expected syllables, without any issue, and as incorrect otherwise. Sentences excluded due to technical issues were not coded and were omitted from the accuracy analysis. A detailed breakdown of excluded and retained sentences by population and condition, along with the reasons for exclusion, is presented in Supplementary Material 1. No outlier removal or a priori cutoffs were applied in our speech analyses, and all retained

sentences, as defined by the criteria above, were included in the subsequent statistical analyses.

Table 2. Description of the speech outcomes used in the current study.

Outcome	**Description**
Accuracy	Binary outcome indicating whether the target sentence was produced as intended (correct/incorrect)
Speech initiation time	Duration between the end of the priming period and the onset of speech production
Speech rate	Syllables per second, including pauses
%V	Proportion of time devoted to vowels, out of the total duration of vocalic and consonantal intervals, expressed in percentage
ΔV	Standard deviation of the durations of vocalic intervals
ΔC	Standard deviation of the durations of consonantal intervals
ΔS	Standard deviation of the durations of syllabic intervals
VarcoV	Rate-normalized variability for vocalic intervals
VarcoC	Rate-normalized variability for consonantal intervals
VarcoS	Rate-normalized variability for syllabic intervals
nPVI-V	Normalized pairwise variability index for vocalic intervals
rPVI-C	Raw pairwise variability index for consonantal intervals
nPVI-S	Normalized pairwise variability index for syllabic intervals

Notes. The formulas used to calculate rhythmic metrics (i.e., the last ten measures in the table) follow Liss et al. (2009) and Lowit et al. (2018).

2.5 Statistical analysis

Statistical analyses were carried out on accuracy, initiation time, speech rate, and rhythmic metrics using mixed models. They were conducted in a Bayesian framework using brms (version 2.22.0) (Bürkner, 2017) together with the Stan MCMC sampler (version 2.37.0) in R (version 4.5.1) (R Core Team, 2021). To test our hypotheses, in all models, fixed effects were population (levels: CTRL [reference], PD, SCA), condition (levels: Silence [reference], Irregular, Regular), musicality score, and age, together with the interactions between condition and population as well as condition and musicality score. Age and musicality score were mean-centered prior to modeling. The random-effects structure included correlated by-subject and by-sentence intercepts and condition slopes, where subject effects were stratified by Population. Note that outcomes were not averaged across sentences prior to modeling. The structure of the models, in brms syntax, is presented in Supplementary Material 2 (SM2).

For accuracy, initiation time and speech rate, univariate models were used. Accuracy was modeled using a Bernoulli likelihood and logit link. Speech initiation time was modeled using an ex-Gaussian likelihood to accommodate right-skewed latencies. Speech rate was modeled using Gaussian likelihood. Regarding rhythmic metrics, to account for residual and random effects correlations, we jointly modeled them using multivariate models under the same fixed- and random-effect structures as above. One model was used for vowels and consonants, i.e., %V, ΔV, VarcoV, ΔC, VarcoC, rPVI-C, and nPVI-V, while a separate model was used for syllables, i.e., ΔS, VarcoS, and nPVI-S. Residual correlations among outcomes were freely estimated. Posterior marginal means and contrasts were obtained with emmeans (version 1.11.1) (Lenth, 2021). In line with our hypotheses, we examined within-population condition contrasts, the interaction between population and condition, and the interaction between musicality score and condition. In all univariate and multivariate models, weakly informative normal priors were used. Models were fitted with 4 chains, 2,000 iterations each (1,000 warm-up) resulting in 4,000 post-warmup draws of the posterior distribution. Convergence and sampling quality were assessed via $\hat{R}$ (all < 1.01), effective sample sizes, and posterior predictive checks. We also assessed divergent transitions, only a few were found for the initiation time model (5 of 4000 – 0.125%). We report posterior estimates as means accompanied by 95% credible intervals (CrI). We also report directional posterior probabilities, $P^+ = Pr(\beta > 0 \mid data)$ and $P^- = Pr(\beta < 0 \mid data)$, $\beta = 0$ indicating the null effect. Evidence is

considered strong when the 95% CrI did not include 0, with a directional posterior probability exceeding 0.975.

If a beneficial effect of rhythmic priming is observed, it is interesting to determine whether this effect is related to the severity of dysarthria, as some individuals may benefit more than others from rhythm. Adopting an exploratory approach, we therefore examined the association between dysarthria severity and the impact of rhythmic priming in the PD and SCA groups, when such an effect was present. To this end, we first computed the mean speech outcome values for each priming condition and each participant, and then derived the contrasts of interest: Regular–Silence, Regular–Irregular, and Irregular–Silence. We subsequently fitted a simple regression model using dysarthria severity scores as the predictor. In addition, as part of our exploratory analyses, we examined the direct association between dysarthria severity and speech outcome measures in PwPD and PwSCA. Dysarthria severity score was used as a fixed effect. The random-effects structure included correlated by-subject and by-sentence intercepts, with subject effects stratified by Population. We applied the same Bayesian modeling approach in our exploratory analyses as in the main analyses. Full model specifications, written in brms syntax, are provided in Supplementary Material 3 (SM3).

3 Results

Table 3 reports descriptive statistics by population (CTRL, PwPD, PwSCA) and condition (Silence, Irregular, Regular) for accuracy, speech initiation time, speech rate, and rhythmic metrics. Figure 2 visualizes condition effects on accuracy, speech initiation time, and speech rate. For the sake of brevity, rhythmic metrics are not shown in this figure. In line with our hypotheses, we focus primarily on the impact of priming condition on speech production, differential responsiveness to condition depending on pathology, and the moderating role of the musicality score on condition effect. We then present the results of the exploratory analyses, i.e., the association between dysarthria severity and our speech outcome measures as well as between dysarthria severity and the rhythmic priming effect. Full Bayesian estimates (posterior means, 95% credible intervals, and directional posterior probabilities) including all the results are provided in the Supplementary Tables presented in SM2 (main analyses) and SM3 (exploratory analyses).

Table 3. Empirical results by population and condition.

Variable	CTRL			PD			SCA		
	S	IR	R	S	IR	R	S	IR	R
Accuracy (proportion)	0.86 (0.11)	0.86 (0.13)	0.89 (0.10)	0.83 (0.17)	0.83 (0.11)	0.85 (0.11)	0.73 (0.18)	0.70 (0.17)	0.73 (0.15)
Initiation time (ms)	1091 (196)	1038 (148)	969 (160)	1047 (295)	1014 (269)	997 (369)	1280 (349)	1223 (300)	1145 (253)
Speech rate (syll/s)	5.04 (0.49)	5.07 (0.65)	5.03 (0.51)	4.96 (0.89)	4.97 (0.75)	4.97 (0.93)	3.94 (0.92)	3.92 (0.84)	3.89 (0.84)
%V	46.39 (2.45)	46.76 (2.72)	45.95 (1.88)	47.72 (2.92)	48.08 (2.44)	47.96 (2.58)	46.66 (2.94)	46.81 (3.68)	46.23 (3.29)
ΔV	44.82 (7.24)	44.47 (5.80)	43.27 (6.90)	46.44 (13.80)	45.54 (8.67)	48.35 (12.58)	61.53 (17.88)	59.70 (14.32)	60.59 (15.79)
ΔC	52.22 (9.65)	50.58 (12.16)	53.25 (9.79)	52.09 (11.17)	52.88 (9.00)	53.22 (13.69)	64.41 (13.07)	65.64 (16.62)	67.89 (13.03)
ΔS	83.32 (14.82)	81.26 (20.07)	84.14 (14.36)	82.61 (18.64)	81.90 (14.88)	82.38 (18.74)	105.83 (24.92)	106.96 (27.11)	111.21 (20.94)
VarcoV	44.47 (4.40)	44.40 (3.98)	43.40 (4.22)	43.79 (5.01)	43.15 (3.40)	44.92 (4.48)	46.77 (4.37)	46.54 (5.22)	47.44 (6.15)
VarcoC	45.46 (5.38)	44.51 (7.19)	46.30 (5.92)	45.85 (5.55)	47.37 (4.68)	46.55 (5.32)	43.77 (4.21)	45.59 (6.24)	46.01 (4.51)
VarcoS	40.70 (5.55)	39.74 (6.26)	40.98 (5.29)	39.37 (4.90)	39.64 (3.75)	39.37 (5.08)	40.33 (4.21)	41.09 (4.97)	42.40 (4.55)
nPVI-V	44.62 (4.08)	43.50 (5.19)	42.72 (4.00)	44.84 (4.98)	44.33 (4.03)	45.26 (4.69)	48.36 (6.09)	46.48 (5.29)	47.37 (6.98)

Table 3. Empirical results by population and condition.

Variable	CTRL			PD			SCA		
	S	IR	R	S	IR	R	S	IR	R
rPVI-C	55.34 (10.29)	55.47 (12.28)	56.42 (8.49)	56.12 (12.47)	56.04 (8.49)	56.54 (14.76)	70.15 (14.59)	72.59 (16.96)	74.77 (12.18)
nPVI-S	44.51 (4.51)	42.49 (5.68)	44.03 (5.28)	43.11 (4.96)	43.31 (4.15)	43.11 (4.16)	46.84 (5.58)	47.08 (5.57)	48.02 (5.46)

Note. Values are mean (SD). CTRL: Controls; SCA: Spinocerebellar ataxia; PD: Parkinson's disease; S: Silence; IR: Irregular; R: Regular.

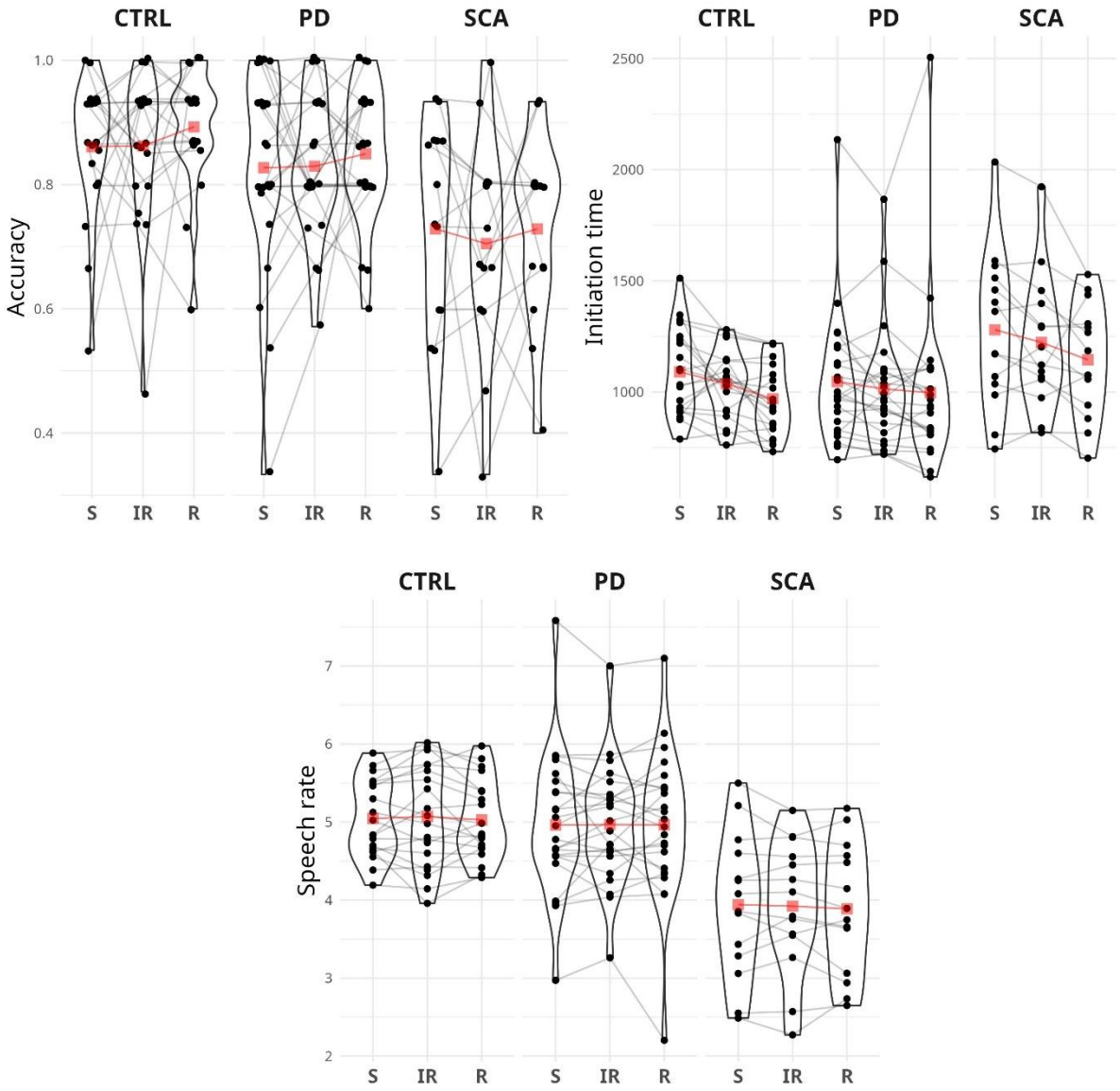

Figure 2. Empirical results by population and condition for accuracy (proportion), initiation time (ms), and speech rate (syll/s). Black dots are subject-level means per condition; gray lines connect the same subject across conditions. Violin plots show the distribution of observations. Red squares are population means, and red segments link population means across conditions. CTRL: Controls; PD: Parkinson's disease; SCA: Spinocerebellar ataxia. S: Silence; IR: Irregular; R: Regular.

3.1 Accuracy

Figure 3 presents posterior estimates of the effect of condition on accuracy as well as the interaction between condition and population. The inspection of the credible intervals for these contrasts revealed no credible effects of condition by population (left) or population and condition interaction (right).

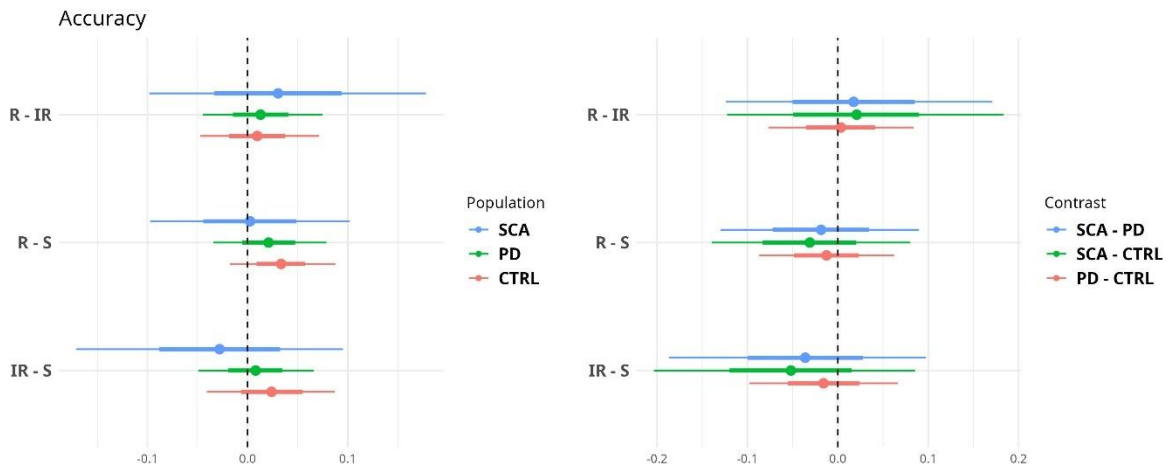

Figure 3. Rhythmic priming effects on accuracy: Posterior within-population condition contrasts (left) and the interaction between condition and population (right). Rows show condition contrasts: Regular vs. Irregular (R–IR), Regular vs. Silence (R–S), and Irregular vs. Silence (IR–S). Dots represent posterior means for each contrast, thick horizontal lines show 66% credible intervals, and thin lines represent 95% credible intervals. Colors encode populations on the left (SCA: Spinocerebellar ataxia; PD: Parkinson's disease; CTRL: Controls) and population comparisons on the right (SCA–PD; SCA–CTRL; PD–CTRL). The dashed vertical line at 0 marks the null effect.

3.2 Speech initiation time

Figure 4 shows posterior estimates of the effect of condition on speech initiation time as well as the interaction between condition and population. A consistent effect of regular primes was observed (Fig. 4, left). For all populations, initiation time was credibly shorter in Regular vs. Silence (R–S): CTRL (−97.24 ms, 95% CrI [−134.35, −61.50]; Pdir = 1.00), PwPD (−47.93 ms, 95% CrI [−73.77, −17.56]; Pdir = 1.00) and PwSCA (−107.92 ms, 95% CrI [−185.53,

−31.55]; Pdir = 1.00). For Regular vs. Irregular (R–IR), a credible effect was present in CTRL (−51.52 ms, 95% CrI [−88.15, −13.92]; Pdir = .99), but not in PwPD or PwSCA. The same pattern was found for Irregular vs. Silence (IR–S) with a credible effect in CTRL (−45.72 ms, 95% CrI [−92.32, −0.27]; Pdir = .98), but not in PwPD or PwSCA. Figure 5 displays subject-level posterior contrasts for R–S and R–IR. The majority of subject means fall below zero across populations, indicating robust shorter initiation times in condition R. Regarding the interaction between condition and population (Figure 4, right), only PwPD vs. CTRL contrast for R–S was credible (49.32 ms, 95% CrI [6.07, 94.77]; Pdir = .99). This indicates that CTRL exhibited a more pronounced reduction in initiation time compared to PwPD.

It is important to note that a common methodological challenge in measuring initiation time involves accurately determining the onset of voicing for voiceless stops. This difficulty may affect empirical results obtained using such linguistic materials (Fargier et al., 2018). While we analyzed the sentences in a very consistent manner, to ensure this issue did not affect our findings, we re-ran the model excluding three sentences with initial voiceless stops. The pattern regarding R–S, R–IR and IR–S remained unchanged; detailed pairwise contrasts among conditions within each population are reported in Supplementary Material 4. This additional analysis indicates that our findings are robust.

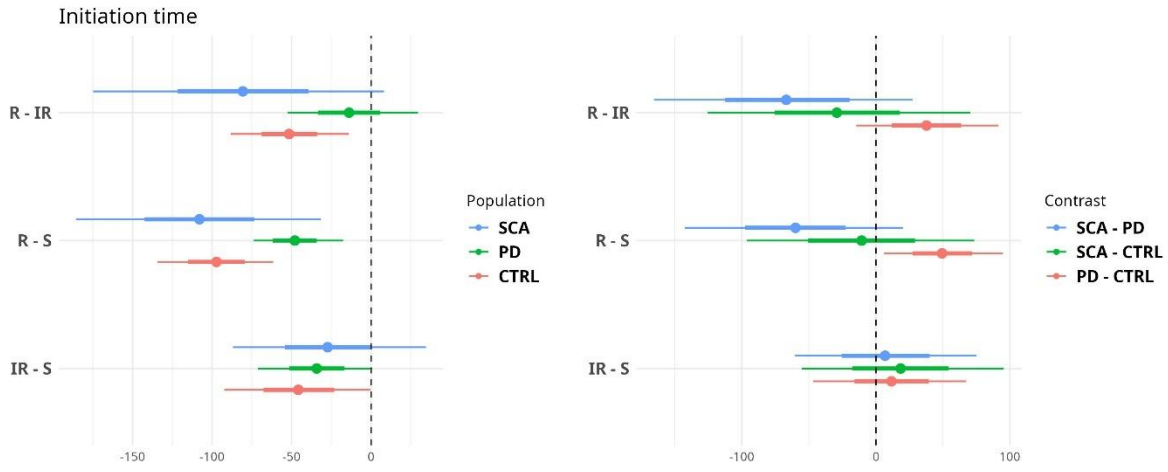

Figure 4. Rhythmic priming effects on speech initiation time. Both panels share the same scale (ms). Legend follows the same conventions as in Figure 3.

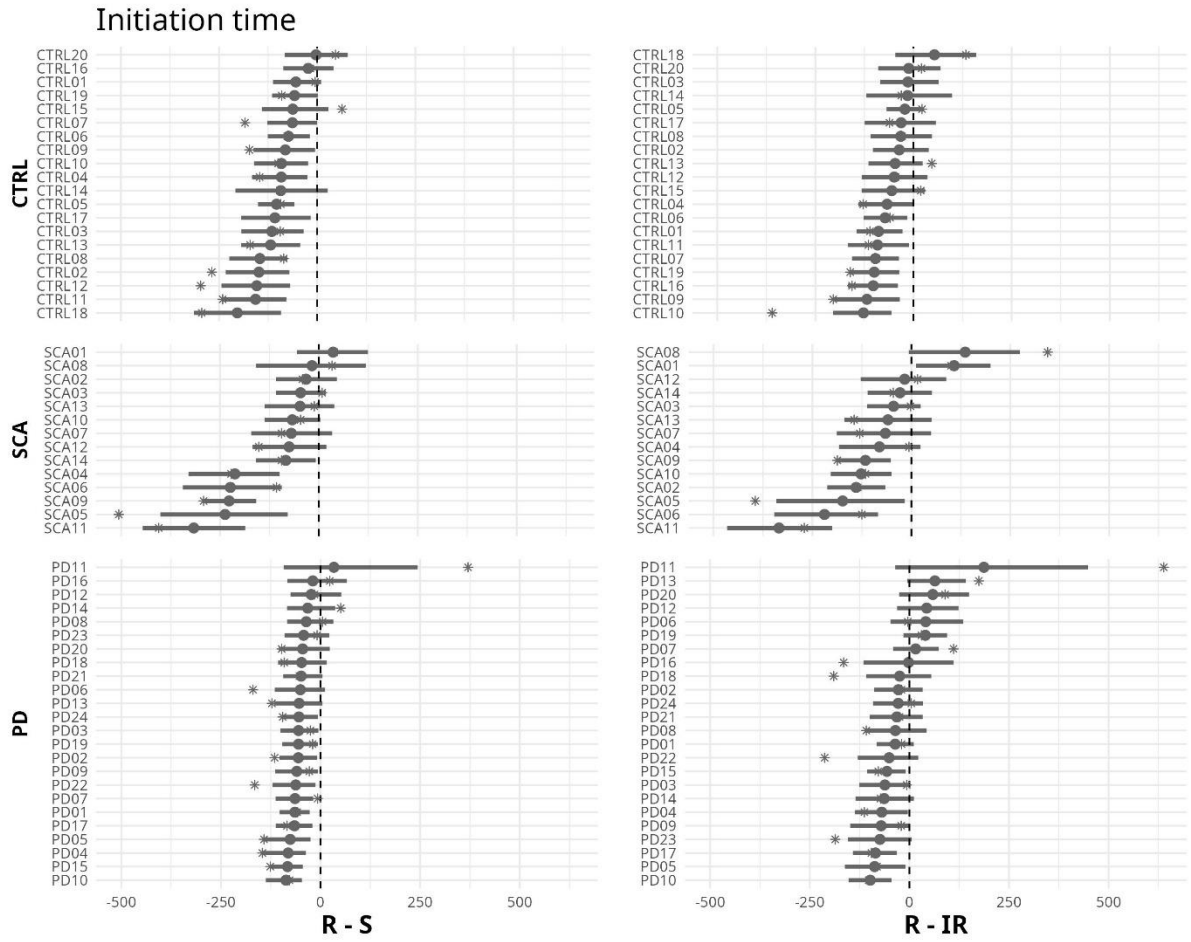

Figure 5. Subject-level posterior estimates for the contrasts Regular vs. Silence (R–S) and Regular vs. Irregular (R–IR) on speech initiation time (ms), shown by population (CTRL: Controls; SCA: Spinocerebellar ataxia; PD: Parkinson's disease). Dots represent posterior means; horizontal bars show 95% credible intervals. Asterisks indicate raw within-subject differences. The dashed vertical line marks 0 (no difference).

3.3 Speech rate

Figure 6 shows the posterior estimates of the effect of condition on speech rate (left) as well as the interaction between condition and population (right). Within-population contrasts showed a credible decrease in speech rate only in PwSCA under Regular vs. Silence conditions (−0.12 syll/s, 95% CrI [−0.23, −0.01]; Pdir = .98). We also examined subject-level posterior contrasts R–S and R–IR (Figure 7). In PwSCA, the visual inspection reveals that R–S estimates show a slight tilt toward negative values for most participants. Regarding the

interaction between condition and population (Figure 6, right), the R–S contrast was larger for PwSCA than to CTRL (−0.15 syll/s, 95% CrI [−0.30, −0.01]; Pdir = .98).

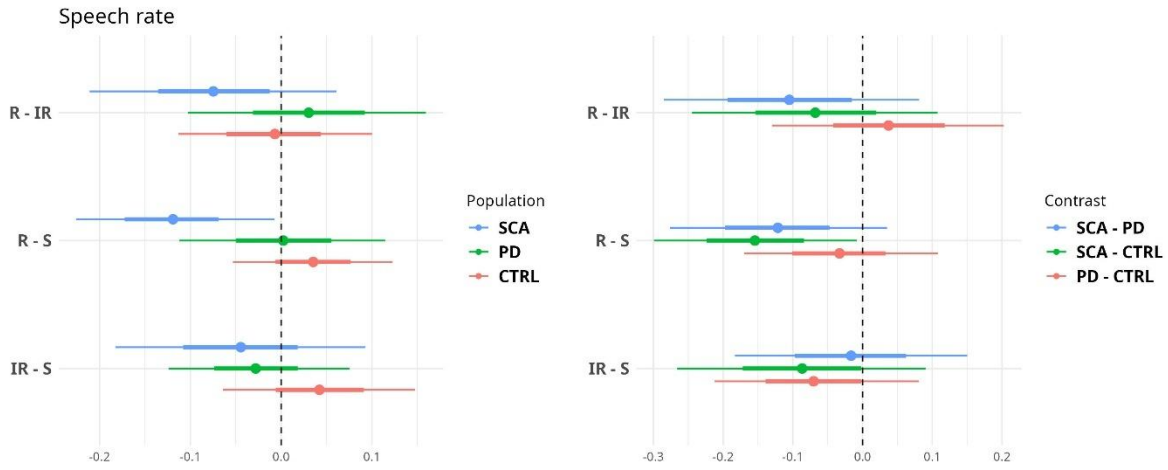

Figure 6. Rhythmic priming effects on speech rate. Both panels share the same scale (syll/s). Legend follows the same conventions as in Figure 3.

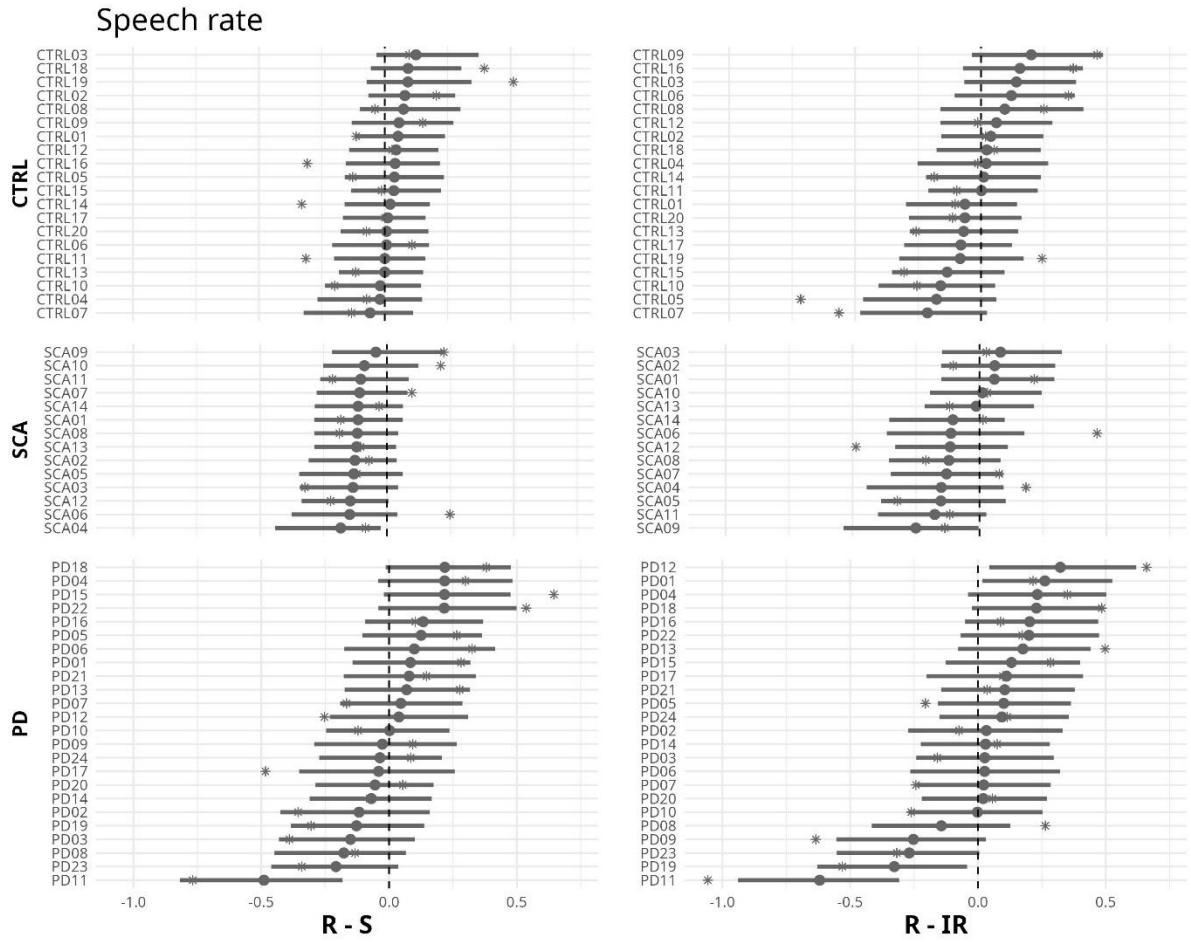

Figure 7. Subject-level posterior estimates for the contrasts Regular vs. Silence (R–S) and Regular vs. Irregular (R–IR) on speech rate (syll/s). Legend follows the same conventions as in Figure 5.

3.4. Rhythmic metrics

For the sake of brevity, we report only the results showing credible effects. For metrics involving vocalic and consonant intervals, the only credible effect was an increase in VarcoC under IR–S for the PwPD (1.78, 95% CrI [0.24, 3.38]; Pdir = .99) (Figure 8, left). Regarding the interaction between condition and population, the only credible effect was for IR–S for PwPD vs. CTRL, showing a larger IR-S contrast in PwPD (2.42, 95% CrI [0.23, 4.54]; Pdir = .99) (Figure 8, right). For metrics involving syllables, credible effects were only observed for ΔS in PwSCA under R–S (5.49, 95% CrI [0.60, 10.37]; Pdir = .99) and R–IR (6.86, 95% CrI [0.94, 12.86]; Pdir = .99) (Figure 9, left). In addition, a small increase in nPVI-S was observed for PwSCA under R–IR (2.32, 95% CrI [0.02, 4.60]; Pdir = .98) (Figure 10, left). Regarding the interaction between

condition and population for these metrics involving syllables, R–S showed a credible effect for the comparison between PwSCA and CTRL in ΔS (6.38, 95% CrI [0.40, 12.67]; Pdir = .98), indicating a larger contrast in ΔS for PwSCA (Figure 9, right).

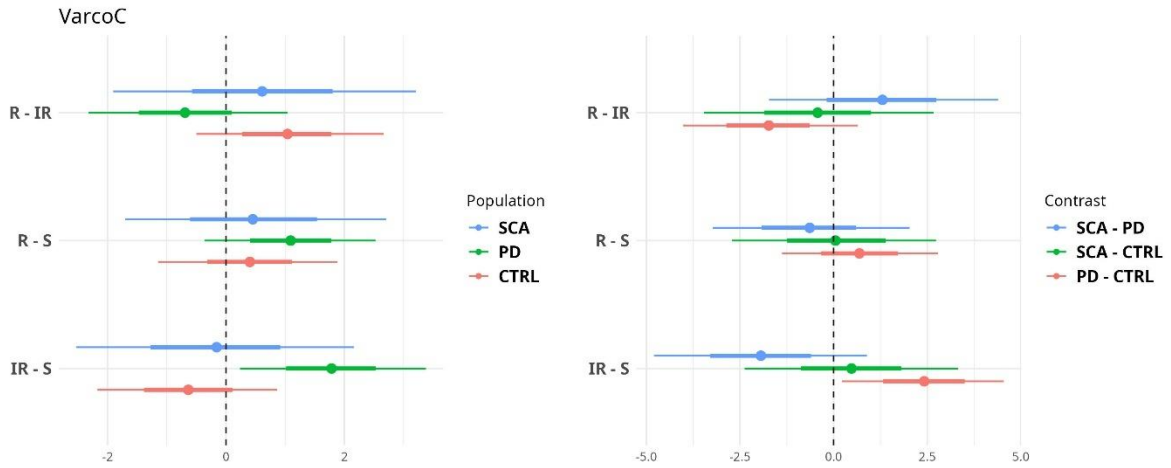

Figure 8. Rhythmic priming effects on VarcoC. Legend follows the same conventions as in Figure 3.

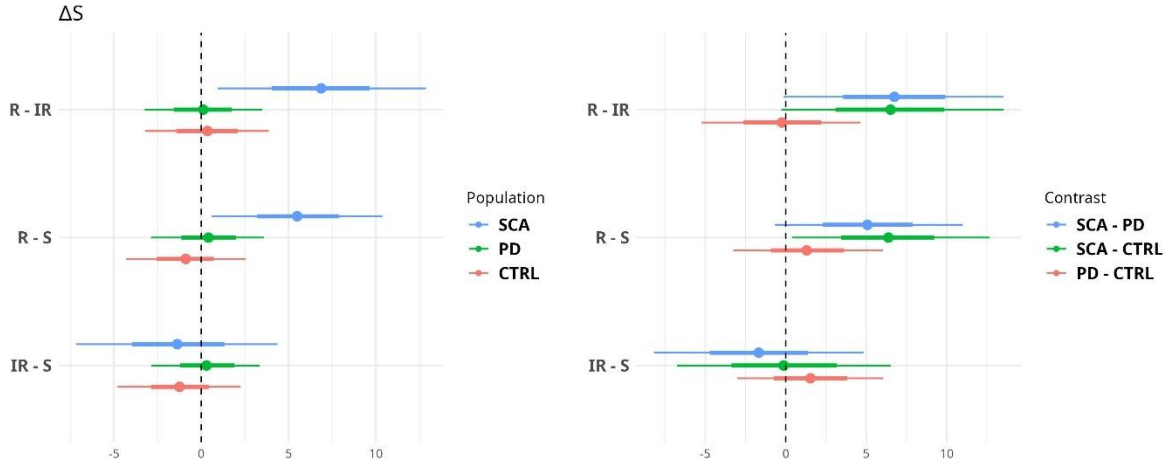

Figure 9. Rhythmic priming effects on ΔS. Legend follows the same conventions as in Figure 3.

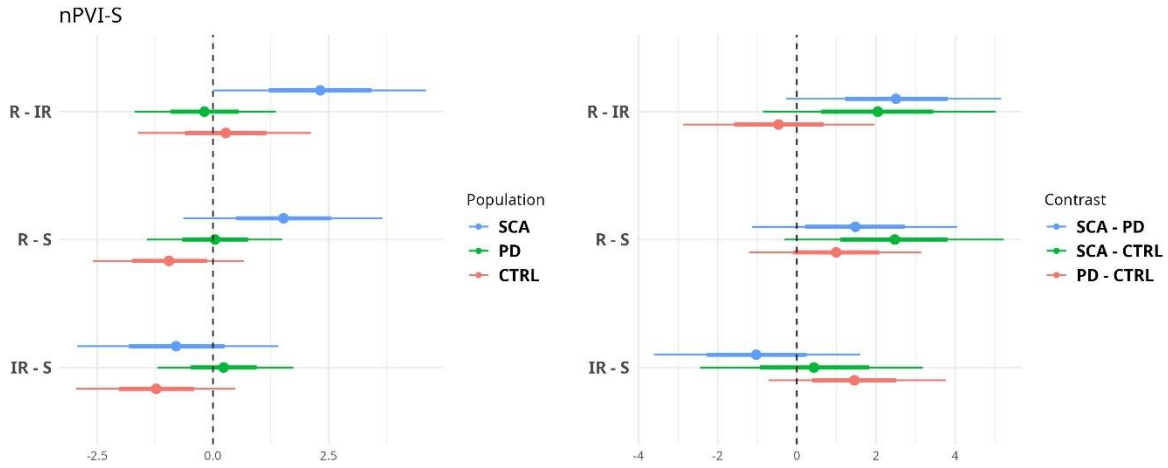

Figure 10. Rhythmic priming effects on nPVI-S. Legend follows the same conventions as in Figure 3.

3.5. Musicality score

We also examined the impact of musicality score on the contrast between conditions across population, for all of the dependent variables, i.e., accuracy, speech initiation time, speech rate, and rhythmic metrics. The only credible association was the interaction between musicality and the R–IR contrast for nPVI-V (−0.28, 95% CrI [−0.53, −0.03]; Pdir = .99). As musicality increases, the R–IR difference becomes more negative, i.e., nPVI-V is relatively lower in the regular condition than in the irregular condition for participants with higher musicality scores (see Supplementary Table 14 in SM2 for all results).

3.6 Dysarthria severity

We also investigated direct associations between dysarthria severity and speech outcomes. In PwPD, credible associations were observed with accuracy and %V: Accuracy was lower and %V was higher in individuals with greater severity (–0.01, 95% CrI [–0.02, 0]; Pdir = .98 and 0.25, 95% CrI [0.03, 0.49]; Pdir = .98, respectively). In PwSCA, credible associations were found with initiation time and speech rate: initiation time was longer and speech rate was slower in individuals with greater severity (23.75, 95% CrI [2.45, 44.91]; Pdir = .98 and –0.13, 95% CrI [–0.21, –0.06]; Pdir = 1.00, respectively). In addition, concerning rhythmic metrics, PwSCA with greater severity showed higher ΔV and ΔC values (1.97, 95% CrI [0.84,

3.07]; Pdir = 1.00 and 1.61, 95% CrI [0.63, 2.62]; Pdir = 1.00, respectively), smaller VarcoC (–0.29, 95% CrI [-0.56, 0]; Pdir = .98), and larger rPVI-C (1.71, 95% CrI [0.67, 2.74]; Pdir = 1.00). No other credible effects were found (see Supplementary Table 6 in SM3). We also examined associations between dysarthria severity and rhythmic priming effects in cases where effects were observed in the main analyses. A credible effect was observed for PwSCA: the rhythmic priming effect on speech initiation time (i.e., smaller initiation time in R compared to S; R–S contrast) increased with greater dysarthria severity (–20.78, 95% CrI [–37.32, –5.32]; Pdir = .99). No other credible effects were observed (see Supplementary Table 7 in SM3).

4 Discussion

The current study aimed to determine whether exposure to rhythmic auditory stimuli facilitates speech production. To investigate this, we employed a rhythmic priming paradigm and examined the impact of regular, irregular, and silence primes on various speech outcomes in PwPD, PwSCA and control participants. The main findings regarding the impact of regular priming are as follows. For accuracy, no credible effect was observed. Regarding initiation time, regular primes markedly reduced speech initiation time across all participant groups. In terms of speech rate, regular primes resulted in a slower speech rate among PwSCA. Concerning rhythmic metrics, regular primes produced larger ΔS and nPVI-S values for PwSCA. No robust evidence was found to support a general moderating role of musical experience on these results. These main findings, along with additional empirical results, are discussed below.

4.1 Impact of rhythmic priming on speech production

As expected, speech initiation time was reduced following a regular prime compared to silence. This effect was observed in all groups. Additionally, speech initiation time after regular primes was credibly shorter than after irregular primes in controls. This effect showed a strong trend in PwSCA but was diminished or absent in PwPD. These results suggest that rhythmic-priming benefits are largely driven by regularity, as a metrically regular prime yielded a robust, cross-group facilitation. The same profile holds at the individual level: most participants initiated faster after a regular prime compared to an irregular prime and silence.

It is noteworthy that the impact of the irregular prime on speech initiation was small and idiosyncratic. To dissociate rhythm-specific benefits from other potential factors in auditory stimuli that may influence the rhythmic priming effect, future work should also include additional control primes, for instance textural sound sequences focusing more on timbre and its associated changes (Canette et al., 2020b). Importantly, the shortening of speech initiation time following regular primes does not reduce accuracy, even in PwSCA, who exhibited lower accuracy compared to other participants. This finding suggests that rhythmic priming facilitates speech initiation without compromising speech precision. The pattern observed here is in line with the findings of Zhang and Zhang (2019) in young adults. This finding can be interpreted based on the notion that a predictable beat enhances temporal preparation, with effects extending beyond the stimulation period, as discussed in the Introduction (Fiveash et al., 2021; Jones, 2019; Kösem et al., 2018; Large & Jones, 1999; Zoefel et al., 2024). Importantly, to our knowledge, our results extend this notion to the speech motor system for the first time. Taken together, these findings support our hypothesis H1.

Regarding between-population differences, in the study of Späth et al. (2022) on rhythmic adaptation, PwSCA did not appear to effectively utilize rhythmic cues presented before a speech production task, whereas PwPD and controls did. In contrast, our study found that PwSCA can indeed benefit from rhythmic cues. To understand this discrepancy, it is noteworthy that some PwSCA in Späth et al.'s study showed rhythmic adaptation effects at the individual level. Additionally, two methodological differences between their study and ours may explain this discrepancy. First, contrary to our approach, Späth et al. used verbal primes (i.e., sentences) rather than non-verbal primes. Second, they examined rhythmic adaptation between a model speaker (prime) and participants, defined as the degree to which participants' produced rhythms converged toward the model speaker's rhythm. To measure this, they used vowel onsets as landmarks to derive p-centers (perceptual centers; subjective moments of sound occurrence, Morton et al., 1976; Rathcke et al., 2021) for both primes and produced sentences. The authors then computed vectors of inter-p-center intervals and quantified rhythmic adaptation as the similarity between the participant's and model speaker's inter-p-center interval patterns. This was the only speech outcome analyzed in their study, making it impossible to determine whether beneficial effects on initiation time were present. Our results should thus be considered complementary rather than contradictory to theirs. Importantly, studies on temporal processing suggest that cerebellar lesions have little

or no impact on rhythm-based processing (e.g., Nozaradan et al., 2017), in contrast to duration-based timing (e.g., Breska & Ivry, 2018; Breska & Ivry, 2021). Despite not explicitly investigating speech production, the observation that PwSCA can initiate speech more rapidly using a regular rhythm aligns with these findings.

Regarding the interaction between condition and population, the beneficial effect of a regular prime was smaller in PwPD than in controls. The differential beneficial effect in PwPD, but not in PwSCA, may reflect the distinct roles of the basal ganglia and cerebellum in temporal processing, as suggested in previous literature (e.g., Breska & Ivry, 2018). This also aligns with evidence that rhythmic abilities are particularly fragile in PD (Puyjarinet et al., 2019). In line with this PD specificity, metrically simple rhythms preferentially recruit the basal ganglia (Grahn & Brett, 2007), and the extended basal ganglia system has been argued to support predictive temporal patterning and hierarchical sequencing (Schwartze & Kotz, 2024). As mentioned in the Introduction, rhythmic auditory stimulation methods have proven effective as a rehabilitation approach for movement disorders in PD (Dalla Bella et al., 2015; Ghai et al., 2018; Koshimori & Thaut, 2018; Nombela et al., 2013; Thaut et al., 2001). Two potential mechanisms underlying these effects include the engagement of the cerebellum as a compensatory system and the stimulation of residual activity within basal ganglia circuits. The attenuated beneficial effects observed in PwPD compared to other participant groups in our study may suggest a greater involvement of the latter mechanism. However, this interpretation remains speculative and needs further investigation in future research.

Regarding the impact of rhythmic priming on speech rate, a between-population effect emerged, with PwSCA showing a slower rate after regular primes, while the speech rate of PwPD and controls was unaffected by prime type. Initially, these findings in PwSCA could suggest a detrimental effect on speech production, but it is more plausible that they reflect a compensatory mechanism. Indeed, PwSCA also exhibited reduced overall accuracy compared to PwPD and controls (see Figures 2 and Supplementary Table 8 in SM2). It is possible that under regular priming, PwSCA attempt to exert greater control over their speech. A regular prime provides an external temporal pattern that PwSCA can lock onto, but exploiting that pattern under a noisy controller tends to involve deliberate slowing to achieve more accurate control. If this is confirmed, it could provide valuable insights for developing therapeutic interventions employed by speech and language pathologists in cases of ataxia. Regarding

PwPD, no such impact of rhythmic priming was observed. The impact of PD on speech rate is heterogeneous across studies, with reports of slower, faster, or more variable rates (Moreau & Pinto, 2019). Illner et al. (2022) reported unchanged speech rates in de novo PwPD during monologue and reading. Our sample, while not de novo, consisted predominantly of PwPD with mild dysarthria. A recent meta-analysis study reported an overall lower —but not reliably different— speech rate in PD compared to controls (Fumel et al., 2024). Therefore, since the speech rate of PwPD was not impaired, it is not possible to draw a definitive conclusion regarding the degree of beneficial impact of rhythmic priming in PwSCA compared to PwPD. Taken together, these findings only partially confirm our hypothesis H2.

Our third hypothesis posited that rhythmic priming would benefit accuracy and rhythmic metrics. Regarding accuracy, no effect of rhythmic priming was observed, despite PwSCA having lower accuracy than both PwPD and controls, indicating potential for improvement. Regarding rhythmic metrics, various metrics were analyzed but very little impact of rhythmic priming was observed. Namely, in PwPD, irregular primes were associated with a small, selective increase in the variability of consonantal intervals, reflected by larger VarcoC under IR–S, and this effect was confined to a single consonant-based metric and was not observed for other measures. In PwSCA, regular primes selectively increased the variability of syllabic intervals, reflected by larger ΔS under R–S and R–IR, and ΔS was overall larger in PwSCA than in controls (see Table 3 and Supplementary Table 8 in SM2). Therefore, larger ΔS values under R–S and R–IR do not move in the direction of normalization. In addition, a small increase in nPVI-S was observed for PwSCA under R–IR; although this effect technically met our credibility criteria, the lower bound of the credible interval was very close to zero, and it should therefore be interpreted with caution. Overall, the impact of rhythmic priming on accuracy and fine-grained segmental and syllabic timing was either limited or absent, contradicting our hypothesis H3.

Regarding the potential moderating role of musical experience, we observed an effect of the musicality score on nPVI-V for the R–IR contrast, with smaller values in R compared to IR. This result suggests that musical experience may influence the beneficial effect of rhythmic priming, consistent with our hypothesis H4. However, nPVI-V was not among the measures affected by regular primes in our study, making it difficult to interpret this musicality effect as evidence supporting its role in the normalization of speech outcomes. Moreover, other

measures that showed a rhythmic priming effect, such as initiation time and speech rate, did not appear to be moderated by the musicality score. In our study, we did not include tasks directly assessing general rhythmic abilities, so the absence of moderation by musicality score may reflect the use of a coarse measure. However, in the study by Späth et al. (2022), no association between auditory rate discrimination and rhythmic adaptation was found. Similarly, Liu and Liu (2024) reported no effect of rhythmic skills in tasks such as rhythm production on the rhythmic priming effect in speech production. Future research could address these inconclusive findings by employing comprehensive assessment protocols such as BAASTA (Battery for the Assessment of Auditory Sensorimotor and Timing Abilities; Dalla Bella et al., 2017), which has also been applied in studies on PD (Puyjarinet et al., 2019) and cerebellar anomalies (Bégel et al., 2022), to further investigate potential links between general rhythmic abilities and rhythmic priming effects on speech production (Rathcke et al., 2021).

Taken together, our findings across different speech outcomes suggest that rhythmic priming mainly influences speech initiation and overall speech rate rather than accuracy or the timing of segments and syllables. We thus suggest that external rhythm mainly affects broad temporal aspects of speech rather than its fine-grained timing control. While this interpretation aligns with previous literature on rhythmic perception-production priming (Jungers & Hupp, 2018; Zhang & Zhang, 2019), other studies have reported divergent findings. In hearing-impaired children, using a comparable paradigm, Cason et al. (2015b) reported a beneficial effect of rhythmic priming on speech production in terms of phonological accuracy. However, their analysis relied on perceptual judgments of correctness for vowels, consonants, syllables, and words, without acoustic analyses. More recently, Liu and Liu (2024) used a similar paradigm with Chinese learners of French and found that rhythmic priming influenced nPVI-V values. Their results showed that the speech of Chinese learners became more similar to that of native French speakers after exposure to rhythmic cues matched with the rhythm of the target French sentences. Using longer musical rhythmic primes, Fiveash et al. (2023) demonstrated that regular primes enhance sentence repetition in children with developmental language disorder. In that study, scoring focused on grammatical accuracy rather than speech features, so the linguistic level of analysis differed from the present work. Differences in methodology and in participants' speech–language profiles may therefore account for discrepancies with our findings.

In addition to our main analyses, which aimed to test our hypotheses concerning the impact of rhythmic priming on speech production, we conducted exploratory analyses on the influence of dysarthria on speech outcomes and on rhythmic priming effects (when present). Among PwPD, credible associations were observed with accuracy and %V. Notably, in their study with English-speaking PwPD and controls, Lowit et al. (2018) identified that the %V metric robustly predicted group membership of participants across various speech tasks. Our observation of a similar pattern in French suggests that %V may serve as a robust marker for dysarthria in PD. In PwSCA, measures related to vowel and consonant variability (ΔV, ΔC, VarcoC, and rPVI-C) were associated with dysarthria severity. This divergence between PwPD and PwSCA likely reflects the fact that hypokinetic and ataxic dysarthria affect different components of speech. Regarding the relationship between dysarthria severity and rhythmic priming effects, PwSCA with more severe dysarthria showed greater benefits in speech initiation time. Interestingly, a comparable pattern in PwPD was reported by Thaut et al. (2001), in which participants spoke in time with a metronome and the assessment concerned intelligibility. If confirmed by future studies, this finding regarding dysarthria severity could have important clinical implications, suggesting that rhythm-based interventions may be particularly beneficial for individuals with more severe dysarthria.

4.2 Underlying neurocognitive mechanisms

As mentioned in the Introduction, the rhythmic priming effect is typically accounted for within the framework of DAT (Large & Jones, 1999; Jones, 2019) and can also be explained within the PRISM framework (Fiveash et al., 2021). These frameworks emphasize the role of rhythm in directing attention toward relevant points in time, thereby making the temporal structure of an event more salient and facilitating predictions about upcoming events. Through neural entrainment, this process enhances performance. Our results align with these frameworks and suggest that similar domain-general mechanisms may be involved in perception-production rhythmic priming and speech production timing. To better understand these mechanisms in relation to speech production, future research on the rhythmic priming effect should clarify the role of neural entrainment by simultaneously analyzing both neural and speech production outcomes. Additionally, exploring the sensorimotor coupling, suggested in PRISM, can provide valuable directions for future work. Notably, motor

engagement has been shown to enhance the impact of rhythmic priming (Cason et al., 2015a; Gindre, 2024), possibly by increasing temporal precision (Berthault et al., 2024). The potential effect of motor engagement on speech production could be investigated using the methodology employed in the current study. If confirmed, this enhancing effect could offer useful insights for clinical applications, particularly since speech-language pathologists already use sensorimotor coupling techniques in practice (Brieussel et al., in press; Daigmorte et al., 2022).

While frameworks such as DAT and PRISM provide essential insights into the rhythmic priming effect and its underlying domain-general mechanisms, they cannot address which specific aspects of speech processing are sensitive to priming. This issue is crucial both for understanding the neurocognitive mechanisms involved and for optimizing clinical applications in speech and language pathology, where identifying the specific components affected is essential. To address this, such effects must be examined within the framework of speech and language models. However, the impact of external rhythms on speech production is not currently incorporated into the models of speech production (e.g., GODIVA/DIVA: Bohland et al., 2010; Guenther, 2016; Tourville & Guenther, 2011; HSFC: Hickok, 2012; ConDialInt: Grandchamp et al., 2019). This holds true for the non-verbal cues investigated in our study, which is not surprising, as most models do not address the link between linguistic and non-linguistic aspects. More surprisingly, these models also fail to account for the impact of external verbal cues on one's own production, i.e., rhythmic adaptation between speakers, as demonstrated by Späth et al. (2022) (see also Aichert et al., 2019). Recently, Guillaume et al. (in revision) developed a new computational model based on Bayesian cue combination frameworks (Ma et al., 2023) to address this issue. This model suggests that rhythmic adaptation in speech production arises from integrating internal cues linked to a speaker's own planning process with external cues derived from perceiving the external rhythm. This model focuses on verbal cues and, in its current form, cannot explain rhythmic priming triggered by non-verbal stimuli. Extending this model to incorporate non-verbal cues, followed by their integration into speech production frameworks, would enhance our understanding of rhythmic perception-production priming effects.

For future modeling work to account for such perception-production effects, the GODIVA/DIVA model includes several elements that make it a promising candidate among

existing models. The model comprises two principal components: the planning loop, responsible for storing forthcoming phonological sequences, and the motor loop, responsible for generating movement commands. According to the GODIVA/DIVA framework, the temporal organization of utterances is governed by the sequential structure buffer of the planning loop (situated in the pre-supplementary motor area), the initiation map of the motor loop (located in the supplementary motor area), and their subcortical projections to the basal ganglia. Within this architecture, the sequential structure buffer encodes the overarching temporal and metrical structure of the utterance, such as syllable frames, at an abstract level that is distinct from its phonological content. Activation of this representation provides input to the initiation map, which, in turn, initiates the corresponding motor programs. Within the GODIVA/DIVA model, speech impairments associated with PD are posited to arise from disruptions in these cortical–basal ganglia circuits (Manes et al., 2024). Although not yet implemented, we suggest that exposure to rhythmic auditory stimuli may activate neural networks involved in speech timing and planning, thereby facilitating speech production through the mechanisms outlined above. This proposal remains speculative and requires future empirical and modeling studies for validation.

4.3 Clinical perspectives

The beneficial impact of rhythmic priming on speech and language processing may open new avenues for therapeutic interventions in speech and language pathology (Schön & Tillmann, 2015). Importantly, music- and rhythm-based interventions are already used in everyday clinical practice by speech and language pathologists (Brieussel et al., in press; Daigmorte et al., 2022), with a marked increase in studies examining their scientific foundations (e.g., Ree et al., 2025; Cancer & Antonietti, 2022). For speech production specifically, Melodic Intonation Therapy (MIT; Sparks & Holland, 1976) is one of the most extensively studied approaches for treating acquired neurogenic communication disorders in adults (Ong et al., 2024). Methods like Rhythmic Speech Cueing® have also shown promise, including in PD (Mainka & Mallien, 2014; Thaut et al., 2001). These methods require patients to synchronize speech with external cues such as a metronome. A key advantage of rhythmic priming is that it preserves the naturalness of speech and segment durations without requiring synchronization with a metronome. Our findings suggest that enhancing global speech timing

control could be a meaningful therapeutic target for disorders affecting speech initiation and global rate rather than fine-grained segmental timing or articulatory impairments. The clinical relevance of this proposal requires validation through interventional studies. Additionally, future studies on clinical applications should also consider existing frameworks and interventions that highlight the emotional and social dimensions of music and rhythm (e.g., Matziorinis & Koelsch, 2022; Overlap, Precision, Emotion, Repetition, and Attention (OPERA) framework; Patel, 2011, 2014). These components warrant further investigation in the context of rhythmic perception-production priming.

Regarding potential clinical applications, it has recently been highlighted that the rhythmic priming effect on sentence comprehension, often tested using grammaticality judgment tasks, may be sensitive to methodological factors (Fiveash et al., 2025). In such paradigms, a musical stimulus typically serves as the prime, followed by a series of sentences (e.g., 32 seconds of music followed by six sentences; Przybylski et al., 2013). It has been suggested that the delay between primes and sentences, as well as the degree of attention directed toward the primes, may attenuate or even abolish the effect (Fiveash et al., 2025). To our knowledge, these factors have not been examined in studies on rhythmic perception-production priming. However, it should be noted that the delay issue does not apply to our design, in which each sentence was immediately preceded by its corresponding prime. In future studies, to better understand the mechanisms underlying the impact of rhythm on speech production and to enable potential therapeutic applications for speech and language production impairments, it is crucial to investigate the methodological characteristics and experimental settings that could optimize the effect.

4.4 Limitations

This study has several limitations that should be acknowledged. Participants showed variability in disease duration and dysarthria severity. This may have affected their responsiveness to rhythmic priming and introduced additional variance. Regarding speech measures, other speech outcomes should also be examined. Accuracy levels, particularly among PwPD, showed a ceiling effect, thereby reducing sensitivity to potential priming effects on this measure. Pause behavior (e.g., pause frequency, duration, and location) was not analyzed in the present study. However, these behaviors may be altered in PD and SCA

(Darling-White & Huber, 2020; Schalling & Hartelius, 2004). Future work should therefore examine whether rhythmic priming modulates pause-related measures and whether such effects are beneficial for PwPD and PwSCA. More broadly, to more thoroughly characterize the locus of the rhythm perception–production priming effect, the impact of rhythmic priming on prosodic aspects of speech warrants testing (Contreras Roa et al., 2023). Accentuation and prominence in relation to the prosodic hierarchical structure in French (Jun & Fougeron, 2000) would be important to examine. This could be done following the approach of Frota et al. (2021) in their study on Portuguese-speaking PwPD. Regarding primes, future work should also orthogonally manipulate regularity and congruence with the sentence structure, an aspect that was not addressed in the current study. In addition, the impact of prime duration should be examined. Two types of primes have been used in previous research without direct comparison: short rhythmic primes, as in our study, and longer musical stimuli, as employed in several investigations of grammatical processing (Fiveash et al., 2025). It would also be valuable to compare linguistic or speech-based primes with non-linguistic rhythmic or musical ones, as these may produce partially distinct transfer effects. Moreover, because this study examined reading aloud, the generalization of these findings to spontaneous speech or extended discourse remains to be explored.

5 Conclusion

Brief nonverbal rhythmic primes robustly facilitated speech initiation in controls, PwPD, and PwSCA. In addition, regular primes selectively slowed the speech rate in PwSCA. No impact on accuracy and only limited effects on rhythmic metrics were observed. These findings suggest that rhythmic priming primarily influences the overall control of speech rather than its fine-grained timing. Our results provide empirical constraints for speech production models, which currently lack explicit mechanisms for external rhythmic input. From a clinical perspective, they offer valuable insights for rhythm-based approaches in speech and language therapy. Future studies should examine the impact of rhythmic priming on additional aspects of speech, such as prominence, and explore the influence of different types of primes on speech production.

Data availability statement

The data that support the findings of this study are available in Zenodo at https://doi.org/10.5281/zenodo.19606497

Acknowledgments

We would like to thank all the participants in this study. We are also grateful to the research assistants for their contribution, to Hélène Lœvenbruck for her comments about speech analyses, and to Charlotte Baey for her help with statistical analyses. This work was supported by the French National Agency for Research (Project ANR-21-CE28-0011) and Association France Parkinson (Fellowship 2025-SHS-0010).

References

Aichert, I., Lehner, K., Falk, S., Späth, M., & Ziegler, W. (2019). Do patients with neurogenic speech sound impairments benefit from auditory priming with a regular metrical pattern? *Journal of Speech, Language, and Hearing Research, 62*(8S), 3104–3118. https://doi.org/10.1044/2019_JSLHR-S-CSMC7-18-0172

Arvaniti, A. (2012). The usefulness of metrics in the quantification of speech rhythm. *Journal of Phonetics, 40*(3), 351–373. https://doi.org/10.1016/j.wocn.2012.02.003

Astésano, C., & Bertrand, R. (2016). Accentuation et niveaux de constituance en français : Enjeux phonologiques et psycholinguistiques. *Langue française, 191*(3), 11–30. https://doi.org/10.3917/lf.191.0011

Atkinson-Clement, C., Letanneux, A., Baille, G., Cuartero, M. C., Véron-Delor, L., Robieux, C., Berthelot, M., Robert, D., Azulay, J. P., Defebvre, L., Ferreira, J., Eusebio, A., Moreau, C., & Pinto, S. (2019). Psychosocial Impact of Dysarthria: The Patient-Reported Outcome as Part of the Clinical Management. *Neuro-degenerative diseases*, *19*(1), 12–21. https://doi.org/10.1159/000499627

Aubanel, V., Bayard, C., Strauss, A., & Schwartz, J.-L. (2020). The Fharvard corpus: A phonemically balanced French sentence resource for audiology and intelligibility research. *Speech Communication, 124*, 68–74. https://doi.org/10.1016/j.specom.2020.07.004

Auzou, P., & Rolland-Monnoury, V. (2019). *Batterie d'évaluation clinique de la dysarthrie*. Ortho Édition.

Brieussel, J., Fumel, J., Ott, L., & Basirat, A. (in press). From neurocognitive frameworks to clinical practice: How music and rhythm are utilized in speech-language pathology. American Journal of Speech-Language Pathology.

Basirat, A. (2025). Language in Parkinson's disease. In L. Cummings (Ed.), *The Oxford handbook of communication disorders in neurodegenerative diseases*. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780198888482.001.0001

Bedoin, N., Brisseau, L., Molinier, P., Roch, D., & Tillmann, B. (2016). Temporally Regular Musical Primes Facilitate Subsequent Syntax Processing in Children with Specific Language Impairment. Frontiers in neuroscience, 10, 245. https://doi-org.ressources-electroniques.univ-lille.fr/10.3389/fnins.2016.00245

Berthault, E., Chen, S., Falk, S., Morillon, B., & Schön, D. (2024). Auditory and motor priming of metric structure improves understanding of degraded speech. *Cognition, 248*, 105793. https://doi.org/10.1016/j.cognition.2024.105793

Bégel, V., Bachrach, A., Dalla Bella, S., Laroche, J., Clément, S., Riquet, A., & Dellacherie, D. (2022). Dance Improves Motor, Cognitive, and Social Skills in Children With Developmental Cerebellar Anomalies. Cerebellum (London, England), 21(2), 264–279. https://doi.org/10.1007/s12311-021-01291-2

Boersma, P. & Weenink, D. (2024). Praat: doing phonetics by computer [Computer program]. Version 6.4.19, retrieved 4 September 2024. https://praat.org

Bohland, J. W., Bullock, D., & Guenther, F. H. (2010). Neural representations and mechanisms for the performance of simple speech sequences. *Journal of Cognitive Neuroscience, 22*(7), 1504–1529. https://doi.org/10.1162/jocn.2009.21306

Breska, A., & Ivry, R. B. (2018). Double dissociation of single-interval and rhythmic temporal prediction in cerebellar degeneration and Parkinson's disease. *Proceedings of the National Academy of Sciences, 115*, 12283–12288. https://doi.org/10.1073/pnas.1810596115

Breska, A., & Ivry, R. B. (2021). The human cerebellum is essential for modulating perceptual sensitivity based on temporal expectations *eLife,* 10:e66743. https://doi.org/10.7554/eLife.66743

Bürkner, P.-C. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software, 80*(1), 1–28. https://doi.org/10.18637/jss.v080.i01

Canette, L.-H., Bedoin, N., Lalitte, P., Bigand, E., & Tillmann, B. (2019). The regularity of rhythmic primes influences syntax processing in adults. *Auditory Perception & Cognition, 2*(3), 163–179. https://doi.org/10.1080/25742442.2020.1752080

Canette, L.-H., Fiveash, A., Krzonowski, J., Corneyllie, A., Lalitte, P., Thompson, D., Trainor, L., Bedoin, N., & Tillmann, B. (2020a). Regular rhythmic primes boost P600 in grammatical error processing in dyslexic adults and matched controls. *Neuropsychologia, 138*. https://doi.org/10.1016/j.neuropsychologia.2019.107324

Canette, L. H., Lalitte, P., Bedoin, N., Pineau, M., Bigand, E., & Tillmann, B. (2020b). Rhythmic and textural musical sequences differently influence syntax and semantic processing in children. *Journal of experimental child psychology*, *191*, 104711. https://doi.org/10.1016/j.jecp.2019.104711

Cason, N., Astésano, C., & Schön, D. (2015a). Bridging music and speech rhythm: Rhythmic priming and audio-motor training affect speech perception. *Acta Psychologica, 155*, 43–50. https://doi.org/10.1016/j.actpsy.2014.12.002

Cason, N., Hidalgo, C., Isoard, F., Roman, S., & Schön, D. (2015b). Rhythmic priming enhances speech production abilities: Evidence from prelingually deaf children. *Neuropsychology, 29*(1), 102–107. https://doi.org/10.1037/neu0000115

Cason, N., & Schön, D. (2012). Rhythmic priming enhances the phonological processing of speech. Neuropsychologia, 50(11), 2652–2658. https://doi.org/10.1016/j.neuropsychologia.2012.07.018

Chern, A., Tillmann, B., Vaughan, C., & Gordon, R. L. (2018). New evidence of a rhythmic priming effect that enhances grammaticality judgments in children. *Journal of Experimental Child Psychology, 173*, 371–379. https://doi.org/10.1016/j.jecp.2018.04.007

Cominu, A. (2021). *Hydrogen* (Version 1.1.1) [Computer software]. Hydrogen-music. http://hydrogen-music.org

Contreras Roa, L., Mairano, P., Moreau, C., & Basirat, A. (2023). A bimodal approach to study the effects of rhythmic priming. In *Proceedings of the 20th International Congress of Phonetic Sciences (ICPhS)* (pp. 3922–3926). Prague, Czech Republic.

Dalla Bella, S., Benoit, C.-E., Farrugia, N., Schwartze, M., & Kotz, S. A. (2015). Effects of musically cued gait training in Parkinson's disease: beyond a motor benefit. *Annals of the New York Academy of Sciences, 1337*, 77–85. https://doi.org/10.1111/nyas.12651

Dalla Bella, S., Farrugia, N., Benoit, C. E., Begel, V., Verga, L., Harding, E., & Kotz, S. A. (2017). BAASTA: Battery for the assessment of auditory sensorimotor and timing abilities. *Behavior Research Methods, 49*(3), 1128–1145. https://doi.org/10.3758/s13428-016-0773-6

Dalrymple-Alford, J. C., MacAskill, M. R., Nakas, C. T., Livingston, L., Graham, C., Crucian, G. P., … Anderson, T. J. (2010). The MoCA: Well-suited screen for cognitive impairment in Parkinson disease. *Neurology, 75*(19), 1717–1725. https://doi.org/10.1212/WNL.0b013e3181fc29c9

Darley, F. L., Aronson, A. E., & Brown, J. R. (1969). Clusters of deviant speech dimensions in the dysarthrias. *Journal of Speech and Hearing Research, 12*(3), 462–496. https://doi.org/10.1044/jshr.1203.462

Darling-White, M., & Huber, J. E. (2020). The impact of Parkinson's disease on breath pauses and their relationship to speech impairment: A longitudinal study. *American journal of speech-language pathology*, *29*(4), 1910-1922. https://doi.org/10.1044/2020_AJSLP-20-00003

Daigmorte, C., Tallet, J., & Astésano, C. (2022). On the foundations of rhythm-based methods in speech therapy. In *Speech Prosody 2022* (pp. 47–51). https://doi.org/10.21437/SpeechProsody.2022-10

Delais, E. (1993). Structure rythmique et prosodique du français : Pour une approche psycho-cognitive. *Linx, 29*(2), 45–71. https://doi.org/10.3406/linx.1993.1272

Duffy, J. R. (2020). *Motor speech disorders: Substrates, differential diagnosis, and management* (4th ed.). Elsevier.

Falk, S., Lanzilotti, C., & Schön, D. (2017). Tuning neural phase entrainment to speech. *Journal of Cognitive Neuroscience, 29*(8), 1378–1389. https://doi.org/10.1162/jocn_a_01136

Fargier, R., Buerki, A., Pinet, S., Alario, F. X., & Laganaro, M. (2018). Word onset phonetic properties and motor artifacts in speech production EEG recordings. *Psychophysiology*, *55*(2), e12982. https://doi.org/10.1111/psyp.12982

Fiveash, A., Bedoin, N., Lalitte, P., & Tillmann, B. (2020). Rhythmic priming of grammaticality judgments in children: Duration matters. *Journal of Experimental Child Psychology, 197*, 104885. https://doi.org/10.1016/j.jecp.2020.104885

Fiveash, A., Bedoin, N., Gordon, R. L., & Tillmann, B. (2021). Processing rhythm in speech and music: Shared mechanisms and implications for developmental speech and language disorders. *Neuropsychology, 35*(8), 771–791. https://doi.org/10.1037/neu0000766

Fiveash, A., Ladányi, E., Camici, J., Chidiac, K., Bush, C. T., Canette, L.-H., Bedoin, N., Gordon, R. L., & Tillmann, B. (2023). Regular rhythmic primes improve sentence repetition in children with developmental language disorder. *NPJ Science of Learning, 8*(1), 23. https://doi.org/10.1038/s41539-023-00170-1

Fiveash, A., Bedoin, N., & Tillmann, B. (2025). Examining methodological influences on the rhythmic priming effect: A commentary on Kim, McLaren, and Lee (2024). *Journal of Experimental Child Psychology, 250*, 106111. https://doi.org/10.1016/j.jecp.2024.106111

Frota, S., Cruz, M., Cardoso, R., Guimarães, I., Ferreira, J. J., Pinto, S., & Vigário, M. (2021). (Dys)Prosody in Parkinson's Disease: Effects of Medication and Disease Duration on Intonation and Prosodic Phrasing. *Brain Sciences*, *11*(8), 1100. https://doi.org/10.3390/brainsci11081100

Fumel, J., Bahuaud, D., Weed, E., Fusaroli, R., & Basirat, A. (2024). A systematic review and Bayesian meta-analysis of acoustic measures of prosody in Parkinson's disease. *Journal of Speech, Language, and Hearing Research, 67*(8), 2548–2564. https://doi.org/10.1044/2024_JSLHR-23-00588

Ghai, S., Ghai, I., Schmitz, G., & Effenberg, A. O. (2018). Effect of rhythmic auditory cueing on parkinsonian gait: A systematic review and meta-analysis. *Scientific reports*, *8*(1), 506. https://doi.org/10.1038/s41598-017-16232-5

Gindre, A. (2024). *Du rythme à la parole : effet d'amorces rythmiques langagières, non langagières et musicales modulées par l'engagement moteur sur le temps d'initiation de la parole* [Doctoral Dissertation, University of Toulouse]. https://theses.hal.science/tel-04931463v1

Goetz, C. G., Tilley, B. C., Shaftman, S. R., Stebbins, G. T., Fahn, S., Martinez-Martin, P., … LaPelle, N. (2008). Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): Scale presentation and clinimetric testing results. *Movement Disorders, 23*(15), 2129–2170. https://doi.org/10.1002/mds.22340

Grahn J. A. (2009). The role of the basal ganglia in beat perception: neuroimaging and neuropsychological investigations. *Annals of the New York Academy of Sciences, 1169*, 35–45. https://doi.org/10.1111/j.1749-6632.2009.04553.x

Grahn, J. A., & Brett, M. (2007). Rhythm and beat perception in motor areas of the brain. *Journal of Cognitive Neuroscience, 19*(5), 893–906. https://doi.org/10.1162/jocn.2007.19.5.893

Grahn, J. A., & Rowe, J. B. (2009). Feeling the beat: Premotor and striatal interactions in musicians and nonmusicians during beat perception. *The Journal of Neuroscience, 29*(23), 7540–7548. https://doi.org/10.1523/JNEUROSCI.2018-08.2009

Grandchamp, R., Rapin, L., Perrone-Bertolotti, M., Pichat, C., Haldin, C., Cousin, E., … Lœvenbruck, H. (2019). The ConDialInt model: Condensation, dialogality, and intentionality dimensions of inner speech within a hierarchical predictive control framework. *Frontiers in Psychology, 10*, 2019. https://doi.org/10.3389/fpsyg.2019.02019

Grube, M., Cooper, F. E., Chinnery, P. F., & Griffiths, T. D. (2010). Dissociation of duration-based and beat-based auditory timing in cerebellar degeneration. *Proceedings of the National Academy of Sciences, 107*(25), 11597–11601. https://doi.org/10.1073/pnas.0910473107

Guenther, F. H. (2016). *Neural control of speech*. MIT Press.

Guillaume, M., Ziegler, W., Timmann, D., Diard, J., & Basirat, A. (in revision). Mechanisms underlying inter-speaker speech adaptation: A computational modeling study with an extension of the Bayesian causal inference model. (Manuscript under revision).

György, D., Saddy, D., Kotz, S. A., & Franck, J. (2024). Rhythmic priming of syntactic processing in Jabberwocky: A short-lived effect. *Language, Cognition and Neuroscience, 39*(7), 939–958. https://doi.org/10.1080/23273798.2024.2361737

Hickok, G. (2012). Computational neuroanatomy of speech production (HSFC). *Nature Reviews Neuroscience, 13*(2), 135–145. https://doi.org/10.1038/nrn3158

Illner, V., Tykalová, T., Novotný, M., Klempíř, J., Dušek, P., & Rusz, J. (2022). Toward automated articulation rate analysis via connected speech in dysarthrias. *Journal of Speech, Language, and Hearing Research, 65*(4), 1386–1401. https://doi.org/10.1044/2021_JSLHR-21-00549

Jones, M. R. (2019). *Time will tell: A theory of dynamic attending*. Oxford University Press.

Jun, S.-A., & Fougeron, C. (2000). A phonological model of French intonation. In A. Botinis (Ed.), *Intonation: Analysis, modelling and technology* (pp. 209–242). Kluwer. https://doi.org/10.1007/978-94-011-4317-2_10

Jungers, M. K., & Hupp, J. M. (2018). Music to my mouth: Evidence of domain-general rate priming in adults and children. *Cognitive Development, 48*, 219–224. https://doi.org/10.1016/j.cogdev.2018.09.001

Kisler, T., Reichel, U., & Schiel, F. (2017). Multilingual processing of speech via web services. *Computer Speech & Language, 45*, 326–347. https://doi.org/10.1016/j.csl.2017.01.005

Koshimori, Y., & Thaut, M. H. (2018). Future perspectives on neural mechanisms underlying rhythm and music-based neurorehabilitation in Parkinson's disease. *Ageing Research Reviews, 47*, 133–139. https://doi.org/10.1016/j.arr.2018.07.001

Kösem, A., Bosker, H. R., Takashima, A., Meyer, A., Jensen, O., & Hagoort, P. (2018). Neural entrainment determines the words we hear. *Current Biology, 28*(18), 2867–2875.e3. https://doi.org/10.1016/j.cub.2018.07.023

Kotz, S. A., & Schwartze, M. (2010). Cortical speech processing unplugged: A timely subcortico-cortical framework. *Trends in Cognitive Sciences, 14*(9), 392–399. https://doi.org/10.1016/j.tics.2010.06.005

Kotz, S. A., & Gunter, T. C. (2015). Can rhythmic auditory cuing remediate language-related deficits in Parkinson's disease? *Annals of the New York Academy of Sciences, 1337*, 62–68. https://doi.org/10.1111/nyas.12657

Kotz, S. A., & Schwartze, M. (2016). Motor-timing and sequencing in speech production: A general-purpose framework. In G. Hickok & S. L. Small (Eds.), *Neurobiology of language* (pp. 717–724). Academic Press. https://doi.org/10.1016/B978-0-12-407794-2.00057-2

Ladányi, E., Lukács, Á., & Gervain, J. (2021). Does rhythmic priming improve grammatical processing in Hungarian-speaking children with and without developmental language disorder? *Developmental Science, 24*(6), e13112. https://doi.org/10.1111/desc.13112

Large, E. W., & Jones, M. R. (1999). The dynamics of attending: How people track time-varying events. *Psychological Review, 106*(1), 119–159. https://doi.org/10.1037/0033-295X.106.1.119

Lenth, R. V. (2021). *emmeans: Estimated marginal means, aka least-squares means* (R package version). https://cran.r-project.org/package=emmeans

Liss, J. M., White, L., Mattys, S. L., Lansford, K., Lotto, A. J., Spitzer, S. M., & Caviness, J. N. (2009). Quantifying speech rhythm abnormalities in the dysarthrias. *Journal of Speech,*

Language, and Hearing Research, 52(5), 1334–1352. https://doi.org/10.1044/1092-4388(2009/08-0208)

Liu, X., & Liu, Y. (2024). Music rhythmic cueing for the production of non-native speech rhythm: Evidence from Chinese learners of French. *Journal of Psycholinguistic Research, 53*(1), 10. https://doi.org/10.1007/s10936-024-10044-1

Lowit, A. (2014). Quantification of rhythm problems in disordered speech: A re-evaluation. *Philosophical Transactions of the Royal Society B: Biological Sciences, 369*(1658), 20130404. https://doi.org/10.1098/rstb.2013.0404

Lowit, A., Marchetti, A., Corson, S., & Kuschmann, A. (2018). Rhythmic performance in hypokinetic dysarthria: Relationship between reading, spontaneous speech and diadochokinetic tasks. *Journal of Communication Disorders, 72*, 26–39. https://doi.org/10.1016/j.jcomdis.2018.02.005

Ma, W. J., Kording, K. P., & Goldreich, D. (2023). *Bayesian models of perception and action: An introduction*. MIT Press.

Manes, J. L., Bullock, L., Meier, A. M., Turner, R. S., Richardson, R. M., & Guenther, F. H. (2024). A neurocomputational view of the effects of Parkinson's disease on speech production. *Frontiers in Human Neuroscience, 18*, 1383714. https://doi.org/10.3389/fnhum.2024.1383714

Matziorinis, A. M., & Koelsch, S. (2022). The promise of music therapy for Alzheimer's disease: A review. *Annals of the New York Academy of Sciences*, *1516*(1), 11–17. https://doi.org/10.1111/nyas.14864

Moreau, C., & Pinto, S. (2019). Misconceptions about speech impairment in Parkinson's disease. *Movement Disorders, 34*(10), 1471–1475. https://doi.org/10.1002/mds.27791

Morton, J., Marcus, S., & Frankish, C. (1976). Perceptual centers (P-centers). *Psychological Review, 83*(5), 405–408. https://doi.org/10.1037/0033-295X.83.5.405

Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., … 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

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

Nombela, C., Hughes, L. E., Owen, A. M., & Grahn, J. A. (2013). Into the groove: Can rhythm influence Parkinson's disease? *Frontiers in Integrative Neuroscience, 7*, 31. https://doi.org/10.1016/j.neubiorev.2013.08.003

Nozaradan, S., Schwartze, M., Obermeier, C., & Kotz, S. A. (2017). Specific contributions of basal ganglia and cerebellum to the neural tracking of rhythm. *Cortex*, *95*, 156-168. https://doi.org/10.1016/j.cortex.2017.08.015

Patel A. D. (2011). Why would Musical Training Benefit the Neural Encoding of Speech? The OPERA Hypothesis. *Frontiers in psychology*, *2*, 142. https://doi.org/10.3389/fpsyg.2011.00142

Patel, A. D. (2014). Can nonlinguistic musical training change the way the brain processes speech? The expanded OPERA hypothesis. *Hearing Research, 308*, 98–108. https://doi.org/10.1016/j.heares.2013.08.011

Pardo, J. S. (2013). Measuring phonetic convergence in speech production. *Frontiers in psychology*, 4, 559. https://doi.org/10.3389/fpsyg.2013.00559

Peeters, G., Giordano, B. L., Susini, P., Misdariis, N., & McAdams, S. (2011). The timbre toolbox: Extracting audio descriptors from musical signals. *The Journal of the Acoustical Society of America*, *130*(5), 2902-2916. https://doi.org/10.1121/1.3642604

de Picciotto, Y., Lithwick Algon, A., Amit, I., Vakil, E., & Saban, W. (2024). Large-scale evidence for the validity of remote MoCA administration among people with cerebellar ataxia. *The Clinical Neuropsychologist, 39*(6), 1634–1650. https://doi.org/10.1080/13854046.2024.2397835

Puyjarinet, F., Bégel, V., Gény, C., Driss, V., Cuartero, M. C., Kotz, S. A., Pinto, S., & Dalla Bella, S. (2019). Heightened orofacial, manual, and gait variability in Parkinson's disease results from a general rhythmic impairment. *NPJ Parkinson's Disease, 5*, 19. https://doi.org/10.1038/s41531-019-0092-6

Przybylski, L., Bedoin, N., Krifi-Papoz, S., Herbillon, V., Roch, D., Léculier, L., Kotz, S. A., & Tillmann, B. (2013). Rhythmic auditory stimulation influences syntactic processing in children

with developmental language disorders. *Neuropsychology, 27*(1), 121–131. https://doi.org/10.1037/a0031277

Ramus, F., Nespor, M., & Mehler, J. (1999). Correlates of linguistic rhythm in the speech signal. *Cognition, 73*(3), 265–292. https://doi.org/10.1016/S0010-0277(99)00058-X

Rathcke, T., Lin, C., Falk, S. & Dalla Bella, S., (2021). Tapping into linguistic rhythm, *Laboratory Phonology 12*(1): 1-32. https://doi.org/10.5334/labphon.248

R Core Team. (2021). *R: A language and environment for statistical computing* [Computer software]. R Foundation for Statistical Computing. https://www.R-project.org/

Schalling, E., & Hartelius, L. (2013). Speech in spinocerebellar ataxia. *Brain and Language, 127*(3), 317–322. https://doi.org/10.1016/j.bandl.2013.10.002

Schmitz-Hübsch, T., du Montcel, S. T., Baliko, L., Berciano, J., Boesch, S., Depondt, C., … Fancellu, R. (2006). Scale for the assessment and rating of ataxia: Development of a new clinical scale. *Neurology, 66*(11), 1717–1720. https://doi.org/10.1212/01.wnl.0000219042.60538.92

Schön, D., & Tillmann, B. (2015). Short- and long-term rhythmic interventions: Perspectives for language rehabilitation. *Annals of the New York Academy of Sciences, 1337*, 32–39. https://doi.org/10.1111/nyas.12635

Schwartze, M., & Kotz, S. A. (2024). Timing patterns in the extended basal ganglia system. In H. Merchant & V. de Lafuente (Eds.), *Neurobiology of interval timing* (Advances in Experimental Medicine and Biology, Vol. 1455). Springer. https://doi.org/10.1007/978-3-031-60183-5_15

Silveri, M. C. (2021). Contribution of the cerebellum and the basal ganglia to language production: Speech, word fluency, and sentence construction—Evidence from pathology. *The Cerebellum, 20*(2), 282–294. https://doi.org/10.1007/s12311-020-01207-6

Späth, M., Aichert, I., Timmann, D., Ceballos-Baumann, A. O., Wagner-Sonntag, E., & Ziegler, W. (2022). The role of the basal ganglia and cerebellum in adaptation to others' speech rate and rhythm: A study of patients with Parkinson's disease and cerebellar degeneration. *Cortex, 157*, 81–98. https://doi.org/10.1016/j.cortex.2022.08.012

Teki, S., Grube, M., Kumar, S., & Griffiths, T. D. (2011). Distinct neural substrates of duration-based and beat-based auditory timing. *Journal of Neuroscience, 31*(10), 3805–3812. https://doi.org/10.1523/JNEUROSCI.5561-10.2011

Teki, S., Grube, M., & Griffiths, T. D. (2012). A unified model of time perception accounts for duration-based and beat-based timing mechanisms. *Frontiers in Integrative Neuroscience, 5*, 90. https://doi.org/10.3389/fnint.2011.00090

Thaut, M. H., Kenyon, G. P., Schauer, M. L., & McIntosh, G. C. (2001). Auditory rhythmicity enhances movement and speech motor control in patients with Parkinson's disease. *Functional Neurology, 16*(2), 163–172.

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

Yesavage, J. A., Brink, T. L., Rose, T. L., Lum, O., Huang, V., Adey, M., & Leirer, V. O. (1983). Development and validation of a geriatric depression screening scale: A preliminary report. *Journal of Psychiatric Research, 17*(1), 37–49. https://doi.org/10.1016/0022-3956(82)90033-4

Walshe, M., & Miller, N. (2011). Living with acquired dysarthria: the speaker's perspective. *Disability and Rehabilitation*, *33*(3), 195–203. https://doi.org/10.3109/09638288.2010.511685

Zhang, N., & Zhang, Q. (2019). Rhythmic pattern facilitates speech production: An ERP study. *Scientific Reports, 9*(1), 1–11. https://doi.org/10.1038/s41598-019-49375-8

Zoefel, B., Abbasi, O., Gross, J., & Kotz, S. A. (2024). Entrainment echoes in the cerebellum. *Proceedings of the National Academy of Sciences, 121*(34), e2411167121. https://doi.org/10.1073/pnas.2411167121

Chapitre 3
Au-delà du rythme

Introduction

Les analyses présentées au chapitre précédent ont permis d'examiner l'effet de l'amorçage rythmique sur la dynamique globale de la production de la parole, à travers le temps d'initiation, le débit et plusieurs métriques rythmiques. Des mesures telles que %V, les indices PVI ou les coefficients de variation apportent des informations utiles sur la distribution temporelle des intervalles vocaliques, consonantiques ou syllabiques. Elles ne suffisent toutefois pas à décrire la manière dont la parole est organisée à l'échelle de l'énoncé, c'est-à-dire la manière dont les variations de durée, de la F0 et les pauses contribuent au phrasé (Arvaniti, 2012, 2020 ; Liss et al., 2009 ; Lowit, 2014).

Cette distinction est importante dans le cadre de la présente thèse. L'enjeu n'est pas seulement de déterminer si un rythme externe peut accélérer ou ralentir la parole, mais aussi d'examiner s'il peut influencer la manière dont celle-ci est organisée. En français, cette organisation repose notamment sur une hiérarchie de constituants prosodiques, au sein de laquelle le groupe accentuel occupe une place centrale. La syllabe finale du groupe accentuel constitue un site privilégié de marquage prosodique, souvent associé à des modifications de durée et de hauteur, tandis que les pauses peuvent contribuer à signaler des frontières prosodiques plus marquées (Astésano & Bertrand, 2016 ; Jun & Fougeron, 2000, 2002).

Cette question prend un relief particulier dans les troubles moteurs de la parole. Dans la MP comme dans l'ASC, les descriptions cliniques classiques rapportent des altérations qui concernent directement ou indirectement la prosodie : réduction des variations mélodiques, ralentissement, allongements anormaux, irrégularité temporelle ou pauses inappropriées (Darley et al., 1969a, 1969b, 1975 ; Duffy, 2020 ; Schalling & Hartelius, 2013). Dans la MP, plusieurs travaux ont plus spécifiquement documenté des modifications de la variabilité de F0, de la dynamique intonative et du phrasé prosodique, ainsi que des altérations liées aux pauses (Frota et al., 2021 ; Lowit & Kuschmann, 2012 ; Thies et al., 2020). Ces éléments rejoignent les résultats de la méta-analyse présentée précédemment dans ce manuscrit, qui

a montré que la diminution de la variabilité de F0 et l'allongement des pauses figuraient parmi les marqueurs acoustiques les plus robustement associés à la MP (Fumel et al., 2024).

Dans cette perspective, le présent chapitre se concentre sur deux dimensions complémentaires de l'organisation prosodique : la variabilité de la F0 et les pauses. La première permet d'examiner la dynamique mélodique globale de l'énoncé, dimension particulièrement pertinente au regard des troubles prosodiques décrits dans la MP et l'ASC. Les secondes permettent d'interroger à la fois la durée des pauses et leur positionnement par rapport aux frontières prosodiques attendues. L'objectif est ainsi de déterminer si l'amorçage rythmique peut modifier certains aspects de l'organisation prosodique interne de l'énoncé.

Méthodes

Les analyses présentées dans ce chapitre reposent sur les productions recueillies dans l'étude décrite au chapitre précédent. Les mêmes critères d'inclusion que ceux appliqués aux analyses du chapitre précédent ont été conservés. Les productions retenues correspondaient donc aux phrases pour lesquelles l'enregistrement était exploitable et dans lesquelles les douze syllabes attendues avaient été réalisées.

Deux dimensions acoustiques ont été analysées. La première concernait la variabilité de la F0, évaluée à partir de l'écart-type de F0, exprimé en demi-tons. Cette mesure visait à caractériser la variabilité mélodique globale de l'énoncé. La seconde concernait les pauses. L'analyse distinguait, d'une part, leur positionnement et, d'autre part, leur durée. Le positionnement a été évalué à partir du nombre de pauses produites aux frontières prosodiques attendues, c'est-à-dire après les syllabes 3, 6 et 9, ainsi qu'à partir du nombre de pauses produites à des positions non attendues. La durée a été évaluée à partir de deux mesures : la durée moyenne des pauses produites aux frontières attendues et la durée moyenne de l'ensemble des pauses produites dans l'énoncé.

Les analyses portaient sur l'impact des trois conditions d'amorçage de l'étude sur ces mesures. Elles visaient à comparer l'effet de ces conditions au sein de chaque population, ainsi qu'à déterminer si la sensibilité aux conditions d'amorçage différait entre les participants

contrôles, les personnes atteintes de la MP et les personnes atteintes d'ASC. Les analyses statistiques ont été conduites dans un cadre bayésien et identique à celui présenté dans le chapitre précédent. En résumé, les modèles ont été ajustés avec le package *brms* sous R, en utilisant le backend *cmdstanr* (Bürkner, 2017). Pour chaque variable dépendante, le modèle incluait les effets fixes de la population, de la condition d'amorçage, du score de musicalité centré, de l'âge centré, ainsi que les interactions entre la population et la condition, et entre le score de musicalité et la condition. La structure générale du modèle était la suivante :

Variable dépendante ~ (Population + Score_Musicalité_centré) × Condition + Age_centré + (1 + Condition |sb| gr(Sujet, by = Population)) + (1 +Condition |st| Phrase)

Les effets aléatoires comprenaient des intercepts et des pentes par condition pour les sujets, estimés séparément selon la population, ainsi que des intercepts et des pentes par condition pour les phrases. Cette structure permettait de tenir compte à la fois de la variabilité interindividuelle, de la variabilité liée aux phrases et de la possibilité que l'effet de l'amorçage varie d'un participant ou d'une phrase à l'autre.

Les variables continues strictement positives, c'est-à-dire la mesure de variabilité de F0 et les mesures de durée des pauses, ont été modélisées à l'aide d'une distribution lognormale. Les variables relatives au nombre de pauses ont été analysées à l'aide de modèles binomiaux avec lien logit, afin d'estimer la probabilité qu'une pause soit produite à une position donnée de la phrase. Pour les pauses attendues, le modèle portait sur les trois frontières prosodiques, situées après les syllabes 3, 6 et 9, et prenait en compte le nombre de ces positions effectivement marquées par une pause. Pour les pauses non attendues, le modèle portait sur les huit autres positions de la phrase, correspondant aux positions syllabiques 1 à 11 à l'exception des positions 3, 6 et 9, et prenait en compte le nombre de pauses produites à ces positions non cibles. Ces deux modèles estimaient donc respectivement la probabilité de produire une pause à une frontière attendue et la probabilité d'ajouter une pause à une position non attendue.

Les effets des conditions d'amorçage ont été examinés à partir des prédictions marginales issues des modèles. Les comparaisons principales portaient sur les contrastes entre conditions au sein de chaque population, en particulier amorce régulière versus absence

d'amorce, amorce irrégulière versus absence d'amorce, et amorce régulière versus amorce irrégulière. Les différences entre populations ont également été examinées pour chaque condition. Enfin, les interactions population × condition ont été interprétées à partir des contrastes de ces effets de condition entre groupes.

Pour chaque contraste, les résultats sont rapportés à partir de la moyenne postérieure, de l'intervalle de crédibilité à 95 % et de la probabilité de direction. Les diagnostics de convergence et de qualité d'ajustement ont été vérifiés à l'aide des diagnostics MCMC usuels et des vérifications prédictives postérieures (pour plus de détails concernant les analyses, voir le chapitre 2 de la deuxième partie du manuscrit).

Dans la suite, nous garderons les abréviations utilisées dans le chapitre 2 de la deuxième partie pour pouvoir plus facilement comparer les résultats des deux chapitres et les discuter. Ainsi, PwPD, PwSCA et CTRL correspondent respectivement aux personnes atteintes de la MP, aux personnes atteintes d'ASC et aux participants contrôles. Les conditions R, IR et S correspondent respectivement à l'amorce régulière, à l'amorce irrégulière et à la condition silence.

Résultats

Par souci de concision, seuls les effets présentant un intervalle de crédibilité à 95 % n'incluant pas 0 et une probabilité de direction supérieure à .975 sont rapportés dans cette section. Les effets ne répondant pas simultanément à ces deux critères ne sont pas interprétés comme crédibles.

Variabilité de F0

Pour la variabilité de F0, le seul effet crédible de condition concernait les PwPD. Dans ce groupe, la variabilité de F0 était plus élevée après une amorce régulière qu'après une amorce irrégulière (R – IR = 0.102 demi-ton, 95% CrI [0.006, 0.199] ; Pdir = .98) (Figure 1). Aucun autre contraste de condition ne présentait d'effet crédible.

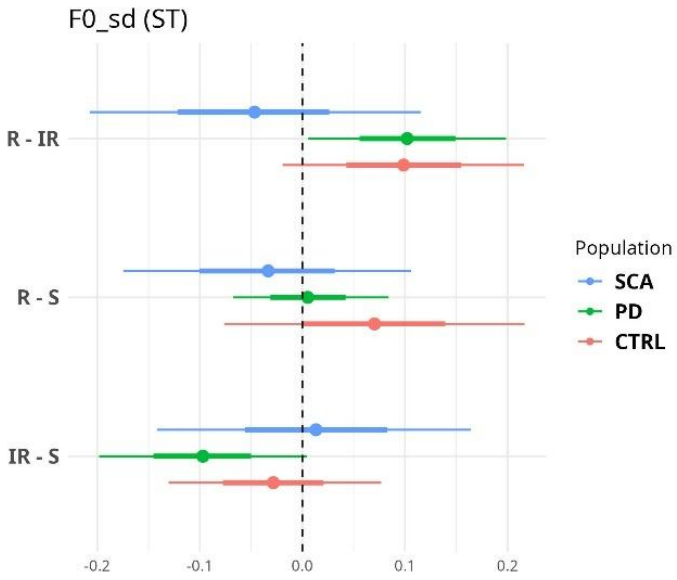

Figure 1. Effets de l'amorçage rythmique sur la variabilité de F0 : contrastes postérieurs intra-population. Les lignes illustrent les comparaisons entre les conditions : Régulier vs. Irrégulier (R–IR), Régulier vs. Silence (R–S) et Irrégulier vs. Silence (IR–S). Les points représentent les moyennes postérieures pour chaque comparaison, les barres horizontales épaisses indiquent les intervalles de crédibilité à 66 % et les barres fines représentent les intervalles de crédibilité à 95 %. Les couleurs codent les populations (SCA : ataxie spinocérébelleuse ; PD : maladie de Parkinson ; CTRL : contrôles). La ligne verticale en pointillés à 0 marque l'effet nul.

Les contrastes entre populations, illustrés sur la Figure 2, mettaient en évidence une variabilité de F0 plus élevée chez les PwSCA que chez les PwPD au niveau global (PwSCA – PwPD = 0.267 demi-ton, 95% CrI [0.011, 0.525] ; Pdir = .98). Cet effet était particulièrement marqué en condition irrégulière (PwSCA – PwPD | IR = 0.353 demi-ton, 95% CrI [0.074, 0.643] ; Pdir = .99). En revanche, aucune différence crédible n'était observée entre les PwSCA et les contrôles, ni entre les PwPD et les contrôles. Aucun contraste d'interaction entre condition et population ne présentait d'effet crédible.

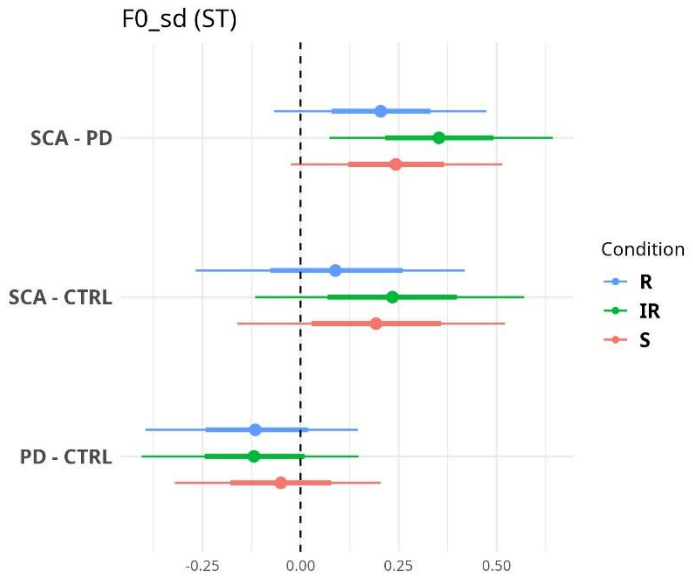

Figure 2. Effets de l'amorçage rythmique sur la variabilité de F0 : contrastes postérieurs intra-condition. Les lignes illustrent les comparaisons entre les populations : PwSCA vs. PwPD (SCA–PD), PwSCA vs. Contrôles (SCA–CTRL) et PwPD vs. Contrôles (PD–CTRL). Les points représentent les moyennes postérieures pour chaque comparaison, les barres horizontales épaisses indiquent les intervalles de crédibilité à 66 % et les barres fines représentent les intervalles de crédibilité à 95 %. Les couleurs codent les conditions (R : Régulier ; IR : irrégulier ; S : Silence). La ligne verticale en pointillés à 0 marque l'effet nul.

Ainsi, les résultats relatifs à la variabilité de F0 suggèrent principalement un effet spécifique de la régularité chez les PwPD, avec une variabilité mélodique plus élevée après amorce régulière qu'après amorce irrégulière. Ils indiquent également une variabilité de F0 plus importante chez les PwSCA que chez les PwPD, en particulier après amorce irrégulière.

Réalisation des pauses

Concernant la réalisation des pauses attendues, les effets crédibles concernaient essentiellement les différences entre populations.

Pour la pause attendue en position 3, les PwSCA présentaient une probabilité plus élevée de réaliser cette pause que les contrôles (PwSCA – CTRL = 0.031, 95% CrI [0.006, 0.080]

; Pdir = .998) et que les PwPD (PwSCA – PwPD = 0.031, 95% CrI [0.005, 0.079] ; Pdir = .997). Cette différence entre PwSCA et contrôles était crédible dans chacune des trois conditions : silence (0.037, 95% CrI [0.003, 0.104] ; Pdir = .99), amorce irrégulière (0.015, 95% CrI [0.001, 0.048] ; Pdir = .99) et amorce régulière (0.042, 95% CrI [0.004, 0.124] ; Pdir = .996). Les PwSCA réalisaient également davantage cette pause que les PwPD en condition silence (0.038, 95% CrI [0.004, 0.105] ; Pdir = .99) et en condition régulière (0.041, 95% CrI [0.002, 0.123] ; Pdir = .99) (Figure 3, haut gauche).

Pour la pause attendue en position 6, aucun effet crédible n'était observé (Figure 3, haut droite).

Pour la pause attendue en position 9, les PwSCA présentaient une probabilité plus élevée de réaliser cette pause que les PwPD au niveau global (PwSCA – PwPD = 0.086, 95% CrI [0.008, 0.230] ; Pdir = .99). Ce contraste était également crédible en condition régulière (PwSCA – PwPD | R = 0.133, 95% CrI [0.020, 0.340] ; Pdir = .997) (Figure 3, bas gauche).

Concernant le nombre total de pauses attendues, les PwSCA produisaient davantage de pauses attendues que les PwPD au niveau global (PwSCA – PwPD = 0.263, 95% CrI [0.002, 0.606] ; Pdir = .98). En condition régulière, les PwSCA réalisaient également davantage de pauses attendues que les contrôles (PwSCA – CTRL | R = 0.325, 95% CrI [0.019, 0.756] ; Pdir = .98) et que les PwPD (PwSCA – PwPD | R = 0.375, 95% CrI [0.052, 0.805] ; Pdir = .99) (Figure 3, bas droite).

Aucun effet crédible de condition n'était observé pour la réalisation des pauses attendues, et aucun contraste d'interaction condition × population ne présentait d'effet crédible. Dans l'ensemble, ces résultats indiquent donc surtout un profil spécifique des PwSCA, caractérisé par une réalisation plus fréquente des pauses attendues, en particulier aux positions 3 et 9, et plus nettement en condition régulière pour le nombre total de pauses attendues réalisées.

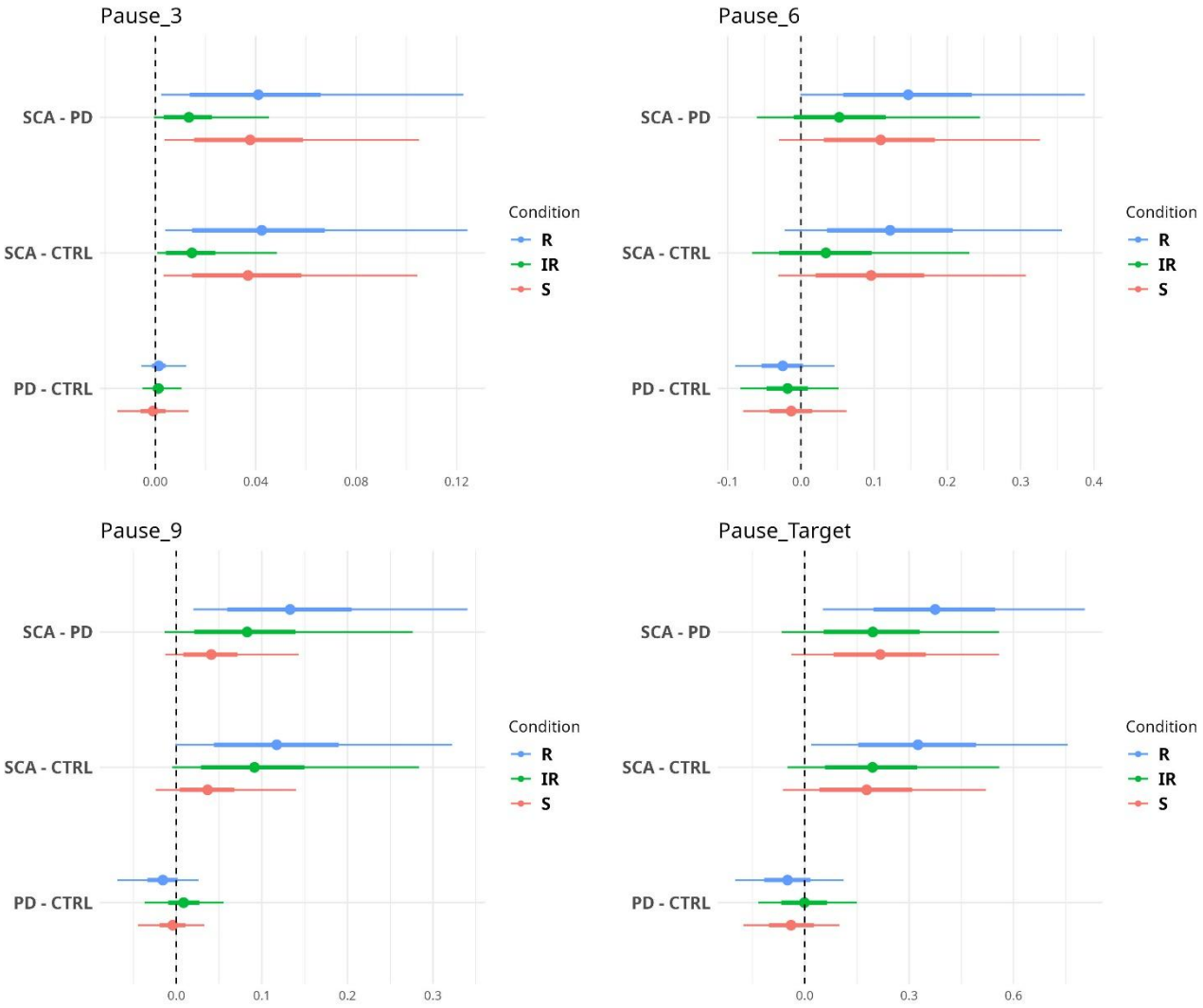

Figure 3. Effets de l'amorçage rythmique sur les pauses attendues : contrastes entre population par condition. Les lignes illustrent les comparaisons entre les populations : PwSCA vs. PwPD (SCA–PD), PwSCA vs. Contrôles (SCA–CTRL) et PwPD vs. Contrôles (PD–CTRL). Les points représentent les moyennes postérieures pour chaque comparaison, les barres horizontales épaisses indiquent les intervalles de crédibilité à 66 % et les barres fines représentent les intervalles de crédibilité à 95 %. Les couleurs codent les conditions (R : Régulier ; IR : irrégulier ; S : Silence). La ligne verticale en pointillés à 0 marque l'effet nul.

Pour les pauses non attendues, aucun effet crédible n'était observé. Aucun contraste de condition, aucun contraste entre populations et aucun contraste d'interaction condition ×

population ne présentait simultanément un intervalle de crédibilité excluant 0 et une probabilité de direction supérieure à 0.975.

Durée des pauses

Pour la durée moyenne de l'ensemble des pauses produites, les PwSCA présentaient des pauses plus longues que les contrôles (PwSCA – CTRL = 0.068 s, 95% CrI [0.016, 0.126] ; Pdir = .99) et que les PwPD (PwSCA – PwPD = 0.059 s, 95% CrI [0.002, 0.121] ; Pdir = .98). Le contraste entre PwSCA et contrôles était crédible dans chacune des trois conditions : silence (0.057 s, 95% CrI [0.001, 0.122] ; Pdir = .98), amorce irrégulière (0.079 s, 95% CrI [0.017, 0.150] ; Pdir = .99) et amorce régulière (0.067 s, 95% CrI [0.010, 0.130] ; Pdir = .99). Les PwSCA présentaient également des pauses plus longues que les PwPD en condition silence (0.064 s, 95% CrI [0.006, 0.130] ; Pdir = .98) (Figure 4, gauche).

Concernant la durée moyenne des pauses attendues, les PwSCA produisaient des pauses attendues plus longues que les contrôles (PwSCA – CTRL = 0.072 s, 95% CrI [0.021, 0.127] ; Pdir = .996) et que les PwPD (PwSCA – PwPD = 0.063 s, 95% CrI [0.009, 0.118] ; Pdir = .99). Le contraste entre PwSCA et contrôles était crédible dans les trois conditions : silence (0.065 s, 95% CrI [0.008, 0.129] ; Pdir = .99), amorce irrégulière (0.082 s, 95% CrI [0.021, 0.147] ; Pdir = .996) et amorce régulière (0.068 s, 95% CrI [0.010, 0.130] ; Pdir = .99). Les PwSCA présentaient également des pauses attendues plus longues que les PwPD en condition silence (0.070 s, 95% CrI [0.012, 0.132] ; Pdir = .99) et en condition irrégulière (0.065 s, 95% CrI [0.002, 0.130] ; Pdir = .98) (Figure 4, droite).

Aucun effet crédible de condition n'était observé pour la durée des pauses, et aucun contraste d'interaction condition × population ne présentait d'effet crédible.

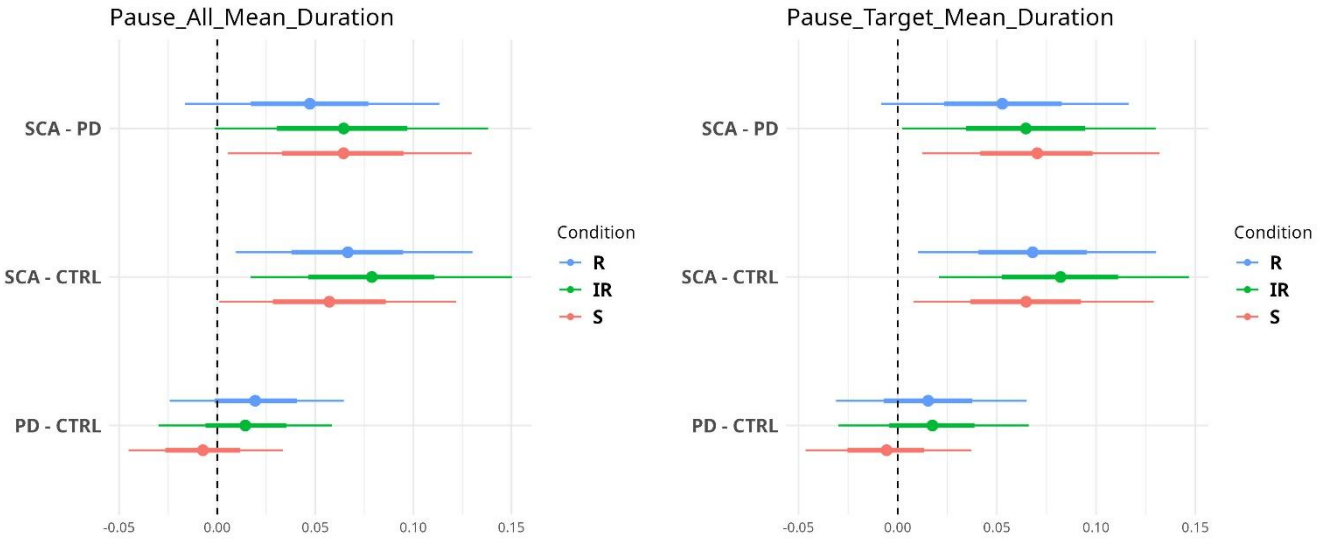

Figure 4. Effets de l'amorçage rythmique sur la durée globale des pauses (gauche) et sur la durée des pauses attendues (droite) : contrastes entre populations par condition. Les lignes illustrent les comparaisons entre les populations : PwSCA vs. PwPD (SCA–PD), PwSCA vs. Contrôles (SCA–CTRL) et PwPD vs. Contrôles (PD–CTRL). Les points représentent les moyennes postérieures pour chaque comparaison, les barres horizontales épaisses indiquent les intervalles de crédibilité à 66 % et les barres fines représentent les intervalles de crédibilité à 95 %. Les couleurs codent les conditions (R : Régulier ; IR : irrégulier ; S : Silence). La ligne verticale en pointillés à 0 marque l'effet nul.

Musicalité

Aucun effet crédible de la musicalité n'était observé pour les variables prosodiques retenues. Les effets principaux de musicalité, les pentes estimées par condition et les contrastes de pente entre conditions ne présentaient pas simultanément un intervalle de crédibilité à 95 % excluant 0 et une probabilité de direction supérieure à .975.

Discussion

L'objectif de ce chapitre était d'examiner si l'amorçage rythmique modulait certaines dimensions prosodiques de la parole produite, au-delà des effets globaux sur l'initiation, le débit et les métriques rythmiques étudiés dans le chapitre précédent. Deux dimensions ont été analysées : la variabilité de F0, comme indice de la dynamique mélodique de l'énoncé, et les pauses, analysées à la fois du point de vue de leur positionnement et de leur durée. Ces

analyses visaient plus particulièrement à déterminer si les amorces régulières pouvaient favoriser certains aspects de l'organisation prosodique chez les PwPD et les PwSCA, si les profils observés différaient entre les PwPD et les PwSCA, et si la musicalité contribuait à moduler ces effets.

Une modulation mélodique limitée mais spécifique

Le résultat le plus directement compatible avec un effet de l'amorçage rythmique concerne la variabilité de F0 chez les PwPD. Dans ce groupe, la variabilité de F0 était plus élevée après une amorce régulière qu'après une amorce irrégulière. Cet effet reste circonscrit : il ne traduit pas une augmentation générale de la variabilité mélodique dans toutes les conditions. Il suggère plutôt que, chez les PwPD, l'opposition entre amorce régulière et amorce irrégulière constitue un contexte particulièrement sensible pour la modulation de la dynamique mélodique.

Ce résultat est intéressant au regard de la littérature sur la dysprosodie parkinsonienne. La MP est fréquemment associée à une réduction des variations mélodiques, à une diminution de la variabilité de F0 et à une parole perçue comme plus monotone ou moins expressive (Duffy, 2020 ; Lowit & Kuschmann, 2012). La méta-analyse présentée précédemment dans ce manuscrit a également identifié la réduction de la variabilité de F0 comme l'un des marqueurs acoustiques les plus robustement associés à la MP (Fumel et al., 2024). Dans ce contexte, le fait que la condition régulière soit associée à une variabilité de F0 plus élevée que la condition irrégulière chez les PwPD peut être interprété comme un indice de modulation mélodique favorable, mais uniquement de manière relative.

Il convient toutefois de rester prudent sur la nature de cet effet. Le contraste crédible oppose la condition régulière à la condition irrégulière, et non la condition régulière à la condition sans amorce. Autrement dit, les résultats ne permettent pas d'affirmer que l'amorce régulière augmente la variabilité de F0 par rapport à une situation de référence neutre. Ils indiquent plutôt que la production mélodique des PwPD diffère selon la structure de l'amorce entendue, avec une variabilité plus élevée après une amorce régulière qu'après une amorce irrégulière. Cette nuance est importante, car elle évite d'interpréter l'effet comme un bénéfice généralisé du rythme régulier. Elle suggère plutôt que les amorces irrégulières et régulières

n'impactent pas de la même manière les mécanismes de préparation ou de maintien de la dynamique prosodique.

Une lecture complémentaire peut être envisagée à partir de la condition irrégulière. Même si le contraste entre la condition irrégulière et la condition silence ne répond pas aux critères de crédibilité retenus dans cette section, l'estimation va dans le sens d'une variabilité de F0 plus faible après amorce irrégulière qu'en condition silence (IR–S = –0.01 s, 95% CrI [–0.019, 0.04] ; Pdir = .97). Cette tendance invite à considérer l'hypothèse selon laquelle l'amorce irrégulière pourrait perturber, chez les PwPD, le maintien d'une dynamique prosodique stable, plutôt que l'amorce régulière ne produise à elle seule un effet de normalisation. Cette interprétation reste prudente, mais elle rejoint les résultats rapportés dans le chapitre 2 de la deuxième partie, où l'amorce irrégulière était associée à une augmentation de la variabilité consonantique mesurée par le VarcoC. Ces éléments suggèrent donc que l'effet observé pourrait tenir autant à une moindre stabilité après amorce irrégulière qu'à un bénéfice propre de l'amorce régulière, hypothèse qui devra être examinée plus directement dans de futurs travaux.

Sur le plan mécanistique, cette interprétation peut être rapprochée des modèles de contrôle moteur de la parole, notamment DIVA et GODIVA, dans lesquels la production repose sur l'interaction entre planification, séquençage, contrôle sensorimoteur et ajustements en ligne (Bohland et al., 2010 ; Tourville & Guenther, 2011). Les ganglions de la base, particulièrement vulnérables dans la MP, sont impliqués dans l'initiation, la sélection et le séquençage des programmes moteurs, ainsi que dans l'intégration de régularités temporelles pertinentes pour l'action (Grahn & Brett, 2009 ; Kotz & Schwartze, 2016 ; Manes et al., 2024). Dans cette perspective, une amorce régulière pourrait fournir un cadre temporel externe susceptible de soutenir certains aspects de la préparation prosodique ou du contrôle mélodique. Cette interprétation doit toutefois rester prudente. Le lien entre la structure temporelle d'une amorce et le contrôle de la F0 n'a pas été directement établi dans la littérature, et les résultats présentés ici demeurent exploratoires. Ils doivent donc être considérés comme une hypothèse de travail, à approfondir dans de futurs travaux plutôt que comme une démonstration mécanistique.

Les résultats intergroupes doivent également être interprétés avec prudence. Les PwSCA présentaient une variabilité de F0 plus élevée que les PwPD, mais aucune différence claire n'était observée entre les PwPD et les contrôles, ni entre les PwSCA et les contrôles. Ce profil ne permet donc pas de conclure à une réduction spécifique de la variabilité mélodique dans la MP dans notre échantillon, ni à une augmentation générale de cette variabilité dans l'ASC. Il suggère plutôt une différence relative entre les deux groupes cliniques, dont l'interprétation doit tenir compte des profils prosodiques habituellement décrits dans ces pathologies. Dans la MP, la réduction de la variabilité de F0 constitue l'un des marqueurs acoustiques les plus robustes de la dysprosodie rapportés dans la littérature (Fumel et al., 2024). Dans les atteintes cérébelleuses, les données disponibles décrivent moins une réduction mélodique systématique qu'une instabilité du contrôle prosodique et phonatoire. Des travaux anciens ont ainsi rapporté des fluctuations anormales de hauteur, une augmentation du jitter ou une variabilité accrue de la F0 dans des tâches de phonation soutenue (Ackermann & Ziegler, 1994 ; Kent et al., 2000). Plus récemment, Cloud et al. (2024) ont également montré, chez des personnes présentant une ataxie cérébelleuse, une variabilité de hauteur plus importante que chez les contrôles dans une tâche de glissando vocal, cette mesure étant associée aux jugements de naturalité de la parole. Ces éléments invitent donc à envisager le contraste observé entre MP et ASC non comme la preuve d'un déficit mélodique spécifique dans notre groupe MP, mais comme un indice compatible avec des signatures prosodiques différenciées selon les pathologies : une tendance à la réduction de la modulation mélodique dans la MP, et un contrôle plus instable des paramètres prosodiques dans les atteintes cérébelleuses.

Les pauses en lien avec la pathologie

Contrairement à la variabilité de la F0, les effets sur les pauses ne relèvent pas principalement de la condition d'amorçage, mais des différences entre populations. Les PwSCA se distinguent par des pauses plus longues que celles des contrôles et des PwPD, que l'on considère l'ensemble des pauses produites ou uniquement les pauses attendues aux frontières prosodiques internes. Ils réalisent également davantage certaines pauses attendues, notamment en positions 3 et 9, ainsi qu'un nombre total plus élevé de pauses attendues dans certaines comparaisons.

Ce profil suggère que les PwSCA ne se caractérisent pas simplement par une plus grande tendance à interrompre la parole, mais par une segmentation plus marquée et plus coûteuse temporellement. Les pauses attendues sont davantage réalisées dans certaines positions, mais elles sont aussi plus longues. Autrement dit, la parole semble davantage découpée, avec des frontières prosodiques plus fortement matérialisées dans le temps. Cette interprétation est cohérente avec les descriptions classiques de la dysarthrie ataxique, qui soulignent l'irrégularité temporelle, la perturbation du débit, l'altération du phrasé et les difficultés de coordination séquentielle (Darley et al., 1969a, 1969b, 1975 ; Duffy, 2020 ; Schalling & Hartelius, 2013 ; Ziegler & Wessel, 1996).

Sur le plan neurofonctionnel, ce résultat peut être mis en lien avec le rôle du cervelet dans la coordination temporelle fine, la prédiction sensorimotrice et l'ajustement précis des séquences motrices. Les atteintes cérébelleuses sont classiquement associées à une difficulté à maintenir la régularité, la fluidité et la précision temporelle des actions, y compris dans la parole (Mariën et al., 2014 ; Teki et al., 2011). Des pauses plus longues chez les PwSCA peuvent ainsi refléter une difficulté accrue à assurer les transitions entre unités prosodiques ou à maintenir une continuité fluide de la production.

Il est également important de noter que toutes les mesures de pauses ne contribuent pas de la même manière au profil observé. Les effets les plus crédibles concernent la durée des pauses, ainsi que la réalisation de certaines pauses attendues. En revanche, les pauses non attendues ajoutées ne montrent pas d'effet crédible. Ce point est important : il suggère que le profil des PwSCA ne correspond pas simplement à une désorganisation globale conduisant à ajouter des pauses partout dans l'énoncé, de façon aléatoire. Il renvoie plutôt à une accentuation temporelle de certaines frontières prosodiques structurales, associée à un allongement des interruptions. Cette distinction est centrale, car elle permet de différencier une parole globalement fragmentée d'une parole dont les frontières attendues deviennent particulièrement coûteuses ou saillantes.

L'absence d'effet crédible de condition sur les pauses invite également à nuancer l'hypothèse d'un effet structurant de l'amorce régulière sur le phrasé. Si l'amorçage rythmique influence certains aspects de la production de la parole, les résultats actuels ne montrent pas qu'il modifie de manière robuste l'organisation des pauses dans l'énoncé. Les différences de pauses semblent ici davantage liées au profil pathologique qu'à la structure de l'amorce. Cela

suggère que, pour cette dimension, les contraintes motrices et temporelles propres à l'ataxie cérébelleuse pèsent davantage que l'effet ponctuel d'un rythme externe bref.

En ce qui concerne les PwPD, les résultats ne suggèrent pas d'effet sur les pauses. Cela peut être lié à la sévérité de la dysarthrie. De plus, la longueur et la complexité de la phrase seraient à prendre en compte, car les contraintes linguistiques pourraient impacter la production des pauses chez les PwPD. En effet, une étude récente a montré que seules les pauses entre les phrases et celles au sein des phrases longues sont significativement plus longues chez les PwPD (Mollaei et al., 2026).

Une dissociation entre modulation mélodique et organisation des pauses

Les résultats suggèrent une dissociation intéressante entre deux dimensions de la prosodie. D'un côté, la variabilité de F0 montre un effet de condition limité mais spécifique chez les PwPD, avec une variabilité plus élevée après amorce régulière qu'après amorce irrégulière. De l'autre, les pauses révèlent surtout un profil propre aux PwSCA, marqué par des durées plus longues et une réalisation plus fréquente de certaines pauses attendues, sans effet robuste de l'amorçage.

Cette dissociation est importante pour l'interprétation générale des résultats de cette thèse. Elle montre que la prosodie ne constitue pas un bloc homogène susceptible d'être uniformément amélioré ou altéré par le rythme. Au contraire, ses composantes peuvent répondre différemment à une amorce rythmique et être affectées différemment selon le profil neuropathologique.

Dans le cas des PwPD, les effets observés sur la F0 doivent être interprétés avec prudence, car notre étude ne met pas en évidence de différence claire entre ce groupe et les contrôles. Ils peuvent toutefois être discutés au regard des difficultés de modulation prosodique classiquement décrites dans la MP. Concernant les PwSCA, les résultats ne montrent pas d'effet crédible de l'amorçage sur les pauses, mais mettent en évidence un profil particulier, caractérisé par des pauses plus longues et par une réalisation plus fréquente de certaines pauses attendues. Ce profil ne suggère donc pas une désorganisation globale du placement des pauses : malgré leur durée accrue, celles-ci apparaissent souvent produites à des positions prosodiquement attendues. Il renvoie davantage à la coordination temporelle, au découpage séquentiel et à la fluidité de la production qu'à une modulation directe par

l'amorce rythmique. Dans l'ASC, l'allongement des pauses pourrait également être discuté au regard des contraintes respiratoires et de coordination pneumo-phonique parfois associées aux dysarthries ataxiques (Kent et al., 2000), sans que les données recueillies ici permettent toutefois de trancher entre une origine respiratoire, motrice ou prosodique. Cette configuration est compatible avec les modèles qui distinguent, sans les opposer de manière rigide, les contributions des ganglions de la base et du cervelet au timing et au contrôle moteur (Grahn & Brett, 2009 ; Kotz & Schwartze, 2016 ; Teki et al., 2011). Les ganglions de la base seraient particulièrement impliqués dans la sélection, l'initiation et le séquençage des actions, tandis que le cervelet contribuerait plus fortement à la précision temporelle, à l'ajustement sensorimoteur et à la coordination fine des transitions. Les résultats obtenus ici ne permettent pas d'inférer directement ces mécanismes, notamment parce que ces modèles portent fréquemment sur l'organisation temporelle plutôt que mélodique de l'action, mais ils s'accordent avec l'idée de profils différenciés de vulnérabilité prosodique.

Limites

Plusieurs limites doivent être soulignées. Premièrement, les analyses de F0 restent globales. Elles permettent d'estimer la variabilité mélodique de l'énoncé, mais ne décrivent pas finement les contours intonatifs locaux, les mouvements de F0 associés aux frontières prosodiques ou les éventuelles différences de réalisation des accents. Une analyse plus fine, par exemple fondée sur des annotations intonatives ou sur des mesures locales de contours, serait nécessaire pour déterminer précisément quelles composantes de la dynamique mélodique sont modulées par l'amorçage.

Deuxièmement, les analyses de pauses reposent sur un matériel très contrôlé, constitué de phrases de douze syllabes organisées en quatre groupes accentuels. Ce contrôle constitue une force méthodologique, car il permet d'identifier des positions attendues de pause et de comparer les participants sur une structure commune. Il limite toutefois la généralisation des résultats à des situations de parole plus spontanée, dans lesquelles le phrasé dépend davantage de choix discursifs, syntaxiques, respiratoires et pragmatiques.

Troisièmement, les effets d'amorçage observés doivent être interprétés dans le contexte d'un amorçage bref et immédiatement antérieur à la production. Les résultats ne

permettent donc pas de conclure sur les effets d'une exposition rythmique plus longue, d'un entraînement répété ou d'une intervention clinique fondée sur le rythme. Ils montrent plutôt qu'un rythme externe bref peut, dans certaines conditions, moduler certaines dimensions spécifiques de la parole, sans produire de réorganisation prosodique globale.

Quatrièmement, les différences interindividuelles liées à la musicalité ont été étudiées ici à l'aide d'un questionnaire. De plus, le niveau musical était globalement faible. Aucun effet crédible n'a été observé. Dans la suite de la thèse, il sera donc important de mieux mesurer les compétences musicales et de distinguer les mesures générales de celles plus directement liées au traitement temporel et au couplage sensorimoteur.

Enfin, les analyses rapportées ici se concentrent volontairement sur les effets crédibles. Cette stratégie permet de limiter les interprétations fragiles, mais elle peut conduire à ne pas discuter certains effets faibles ou incertains qui pourraient devenir informatifs dans des échantillons plus larges ou dans des designs spécifiquement conçus pour tester ces dimensions. Les conclusions doivent donc être comprises comme prudentes et centrées sur les effets les plus robustes.

Conclusion

Pour conclure, les analyses sur la variabilité de la F0 et les pauses apportent un éclairage complémentaire aux résultats du chapitre précédent. Elles ne confirment pas l'hypothèse d'un bénéfice global des amorces régulières, mais elles montrent que certaines dimensions de la parole peuvent être modulées de manière sélective. Chez les PwPD, l'effet le plus crédible concerne la variabilité de F0, plus élevée après amorce régulière qu'après amorce irrégulière. Chez les PwSCA, les résultats mettent surtout en évidence un profil distinct, caractérisé par des pauses plus longues et par une réalisation plus fréquente de certaines pauses attendues.

Ces résultats soutiennent donc une lecture multidimensionnelle des effets du rythme sur la parole. L'amorçage rythmique ne semble pas agir comme un levier uniforme

d’amélioration, mais comme un contexte susceptible d’interagir avec certaines composantes de la production, selon la variable étudiée et le profil neuropathologique.

Chapitre 4
Du rythme perçu à la parole produite : synchronisation sensorimotrice et amorçage rythmique

Introduction

Les chapitres précédents ont montré qu'un rythme auditif présenté avant la parole peut moduler certains aspects de la production orale. Ces résultats s'inscrivent dans une littérature plus large suggérant que l'exposition préalable à une structure rythmique peut faciliter différents traitements langagiers, notamment phonologiques, syntaxiques ou perceptifs, selon les tâches et les populations étudiées (Bedoin et al., 2016 ; Canette et al., 2019 ; Canette et al., 2020a ; Cason & Schön, 2012 ; Cason et al., 2015 ; Fiveash et al., 2020 ; Kotz & Gunter, 2015 ; Ladányi et al., 2021 ; Przybylski et al., 2013). Dans les paradigmes d'amorçage rythmique, l'hypothèse générale est qu'un rythme entendu avant un stimulus langagier peut mettre le système dans un état temporel favorable, en renforçant l'anticipation des événements à venir et en facilitant l'organisation du traitement subséquent.

Cette hypothèse repose en partie sur l'idée que le langage, et plus particulièrement la parole, est un phénomène profondément temporel. Produire une phrase implique non seulement de sélectionner et d'organiser des unités linguistiques, mais aussi de les inscrire dans une dynamique temporelle précise : déclencher l'énoncé au bon moment, organiser les syllabes, répartir les proéminences, ajuster le débit et structurer les groupes prosodiques. La parole ne peut donc pas être pensée uniquement comme une séquence d'unités segmentales ; elle repose également sur des régularités rythmiques et prosodiques qui contribuent à sa planification comme à sa réalisation. Dans cette perspective, une amorce rythmique pourrait influencer la production de la parole en fournissant des repères temporels susceptibles de préparer l'entrée dans la parole ou d'orienter certains aspects de son organisation temporelle.

Comme présenté dans la première partie de ce manuscrit, plusieurs cadres théoriques permettent de comprendre ce type d'effet. La DAT propose qu'un rythme externe puisse aligner les fluctuations attentionnelles internes sur les moments temporellement attendus, facilitant ainsi le traitement d'événements prévisibles (Jones & Boltz, 1989 ; Large & Jones, 1999 ; Jones, 2019). Cette approche permet d'envisager l'amorce rythmique comme un contexte temporel préparatoire : en rendant certaines régularités plus saillantes, elle pourrait faciliter la transition entre perception du rythme et production de parole. Dans une perspective complémentaire, le cadre PRISM propose que les liens entre rythme musical, parole et langage reposent sur trois mécanismes partiellement partagés : un traitement auditif précis des indices temporels, une synchronisation aux régularités externes et un couplage sensorimoteur permettant d'articuler perception et action (Fiveash et al., 2021). Même lorsque l'entraînement neuronal n'est pas directement étudié, ce cadre reste particulièrement utile pour penser la manière dont une structure rythmique perçue peut être convertie en ressource pour une action vocale ultérieure.

La composante sensorimotrice est ici centrale. Si un rythme externe facilite la parole, ce n'est probablement pas seulement parce qu'il est entendu ou reconnu comme régulier, mais parce qu'il peut être intégré dans un système perception–action capable d'exploiter cette régularité pour préparer le mouvement. La production de parole est elle-même une activité motrice finement temporalisée, reposant sur la coordination d'effecteurs respiratoires, laryngés et articulatoires. De ce point de vue, les compétences de synchronisation rythmique pourraient constituer un facteur important de variabilité interindividuelle. Des travaux sur la synchronisation auditivo-motrice montrent en effet que les individus diffèrent dans leur capacité à se synchroniser à des stimuli rythmiques, à maintenir un tempo ou à extraire une pulsation régulière (Dalla Bella et al., 2017b ; Fujii & Schlaug, 2013). Ces différences pourraient moduler la manière dont une amorce rythmique est utilisée pour préparer la parole.

Cette question est d'autant plus importante que les effets de l'amorçage rythmique ne semblent pas toujours homogènes. La régularité du stimulus est souvent considérée comme un facteur clé : une amorce régulière fournirait un échafaudage temporel plus prédictible qu'une amorce irrégulière. Cependant, la régularité ne suffit probablement pas à expliquer l'ensemble des effets observés. Plusieurs études suggèrent que la relation entre la structure

de l'amorce et celle du matériel langagier cible peut également jouer un rôle. Dans des tâches perceptives ou linguistiques, les bénéfices de l'amorçage peuvent dépendre de la congruence ou de la compatibilité entre la structure métrique de l'amorce et celle de la cible (Cason & Schön, 2012 ; Cason et al., 2015). Il est donc nécessaire de déterminer dans quelles conditions précises une structure rythmique préalable facilite effectivement la production.

Dans le domaine de la production orale, cette question reste encore relativement peu documentée. La majorité des travaux sur l'amorçage rythmique se sont intéressés à la perception, à la compréhension ou au traitement linguistique, alors que la production de parole implique des contraintes spécifiques de planification et de mise en mouvement. Comme vu précédemment, quelques études suggèrent néanmoins qu'une exposition rythmique préalable peut influencer la production, notamment le temps d'initiation ou certains paramètres temporels de l'énoncé produit (Jungers & Hupp, 2018 ; Zhang & Zhang, 2019). Ces résultats invitent à distinguer plusieurs niveaux possibles d'effet. Une amorce peut agir sur la préparation de la réponse, mesurée par le temps d'initiation, sans nécessairement modifier le déroulement temporel de l'énoncé une fois la parole engagée. À l'inverse, elle pourrait aussi influencer le débit ou la structuration temporelle globale de la phrase produite. Cette distinction est importante, car elle permet de préciser à quel niveau de la production l'amorçage rythmique exerce son effet.

Le présent chapitre s'inscrit dans cette problématique. Il examine, chez des adultes francophones sans pathologie neurologique, l'effet d'amorces rythmiques non verbales sur deux marqueurs de production : le temps d'initiation de la parole et le débit de parole. Le temps d'initiation renseigne sur la transition entre l'écoute de l'amorce et le déclenchement de la parole ; le débit de parole renseigne davantage sur la dynamique temporelle globale de l'énoncé une fois la production commencée. Le plan expérimental croise la régularité de l'amorce, sa structure rythmique et la structure prosodique de la phrase cible. Il permet ainsi d'examiner non seulement si les amorces régulières facilitent la production par rapport aux amorces irrégulières, mais aussi si cet effet dépend de la compatibilité entre la structure de l'amorce et celle de la phrase à produire.

Un second objectif est d'examiner le rôle des différences individuelles rythmiques et musicales. Dans la logique du cadre PRISM, les effets d'une amorce rythmique devraient dépendre, au moins en partie, de la capacité du participant à extraire une structure temporelle

et à la convertir en ressource pour l'action. Les tâches de tapping permettent ici d'approcher différentes dimensions de cette capacité, notamment la consistance de synchronisation, le maintien d'un tempo et la sensibilité à l'organisation métrique. Les scores musicaux permettent quant à eux de prendre en compte le rapport plus général des participants à la musique et à la pratique musicale. Ces mesures sont utilisées de manière exploratoire afin de déterminer si les compétences rythmiques et musicales sont associées aux effets observés sur la parole, en particulier aux variations du temps d'initiation et du débit selon les conditions d'amorçage.

Ainsi, ce chapitre vise à préciser les conditions dans lesquelles un rythme perçu peut influencer la parole produite. L'hypothèse principale est que les amorces régulières devraient favoriser une initiation plus rapide de la parole par rapport aux amorces irrégulières, en fournissant des repères temporels plus prédictibles. Une hypothèse complémentaire est que cet effet pourrait être modulé par la structure de l'amorce et par sa compatibilité avec la structure prosodique de la phrase cible. Enfin, dans une perspective de différences individuelles, nous examinons si les participants présentant de meilleures compétences de synchronisation rythmique ou une expérience musicale plus importante montrent une sensibilité accrue aux effets de l'amorçage. Cette approche permet de replacer les effets comportementaux observés dans une réflexion plus large sur les relations entre perception du rythme, couplage sensorimoteur et préparation de la parole. Il convient enfin de préciser que cette étude comportait également un volet EEG. L'activité cérébrale des participants a été enregistrée pendant le protocole, afin de permettre ultérieurement l'analyse du suivi neurophysiologique des amorces rythmiques, notamment à travers des mesures de cohérence cerveau–stimulus (Kösem et al., 2018). Ces analyses ne sont pas présentées dans ce chapitre, qui se concentre sur les effets comportementaux de l'amorçage rythmique et sur leurs liens avec les compétences rythmiques individuelles. Elles constituent néanmoins un prolongement important de ce travail, car elles permettront de mieux préciser les mécanismes par lesquels une structure rythmique perçue peut préparer ou moduler la production de la parole.

Méthodes

Participants

Trente locuteurs natifs du français ont été inclus dans les analyses présentées dans ce chapitre (17 femmes ; âge moyen = 21,9 ans, ET = 2,7). La durée moyenne de scolarité était de 15,3 années (ET = 1,6). Tous étaient droitiers, comme vérifié à l'aide de l'*Edinburgh Handedness Inventory*, et ne rapportaient aucun antécédent neurologique, langagier ou auditif. Leur vision était normale ou corrigée. Une compensation financière a été proposée pour leur participation. L'étude a été menée conformément aux principes de la Déclaration d'Helsinki et a reçu l'approbation du comité d'éthique local. Tous les participants ont signé un formulaire de consentement éclairé après avoir reçu une information détaillée sur les objectifs et le déroulement du protocole. Un score cognitif a été recueilli à l'aide du *Montreal Cognitive Assessment* (MoCA ; Nasreddine et al., 2005). Une audiométrie tonale liminaire a également été réalisée afin de confirmer l'absence de trouble auditif. Les seuils auditifs ont été mesurés aux fréquences de 500, 1000, 2000 et 4000 Hz dans chaque oreille, puis moyennés pour chaque participant. Le seuil auditif moyen de l'échantillon retenu était de 5,37 dB (ET = 3,6). Les participants complétaient en outre un questionnaire portant sur leur formation et leur pratique musicales, à partir duquel était calculé un score de musicalité. La sensibilité aux récompenses musicales était évaluée au moyen de la version française du *Barcelona Music Reward Questionnaire* (Saliba et al., 2016).

Matériel

Le matériel linguistique était constitué de 256 phrases déclaratives en français. Parmi elles, 45 provenaient de l'étude présentée dans les chapitres 2 et 3 de la deuxième partie et 211 ont été construites selon la même méthodologie afin d'étendre le corpus expérimental. Les phrases ont été élaborées de manière à contrôler leur organisation prosodique. Dans le présent matériel, toutes les phrases comportaient 12 syllabes, mais variaient quant à leur organisation en trois (S3) ou quatre (S4) groupes accentuels, ce qui permettait de manipuler la structure prosodique sans modifier la longueur syllabique globale. Pour les phrases S4, l'organisation temporelle attendue correspondait à quatre groupes successifs de trois syllabes, soit une alternance de type *xxX xxX xxX xxX* (x : syllabe non proéminente ; X : syllabe

proéminente). Pour les phrases S3, les 12 syllabes étaient regroupées en trois AP, soit xxxX xxxX xxxX, produisant une distribution différente des syllabes proéminentes à l'échelle de la phrase, tout en conservant un même nombre total de syllabes. Conformément à certains modèles de la prosodie du français, ces AP pouvaient également être regroupés en constituants prosodiques supérieurs, tels que les syntagmes intermédiaires (intermediate phrases, ip) (Jun & Fougeron, 2002 ; Delais-Roussarie et al., 2015).

Les amorces rythmiques ont été construites également de façon similaire à celles utilisées dans les chapitres 2 et 3. Les mêmes sons percussifs ont été utilisés. À partir de ces unités, quatre amorces ont été construites : deux amorces régulières et deux amorces irrégulières, chacune comportant 12 événements. Les deux amorces régulières ont été conçues pour refléter la structure temporelle globale des deux types de phrases. L'une correspondait à une organisation à quatre accents, compatible avec la structure attendue des phrases S4, tandis que l'autre correspondait à une organisation à trois accents, compatible avec celle des phrases S3. Les deux amorces irrégulières ont été générées par randomisation de la position des événements accentués et non accentués, tout en conservant le même nombre total d'événements et le même nombre d'accents que les amorces régulières correspondantes, soit trois pour l'une et quatre pour l'autre. Plusieurs variantes ont été produites pour chaque amorce irrégulière. Une amorce irrégulière de chaque type a ensuite été sélectionnée sur des critères perceptifs auprès de sept personnes. La durée totale des amorces était de 2,4 secondes.

Procédure

Chaque session expérimentale était conduite individuellement et durait environ 2 h 30. Elle débutait dans une salle calme, par la lecture de la note d'information, d'éventuelles explications complémentaires et la signature du formulaire de consentement. Les participants complétaient ensuite un questionnaire de sélection, puis passaient les évaluations préalables prévues par le protocole : Montreal Cognitive Assessment (MoCA), audiométrie tonale liminaire, questionnaire sur la pratique et la formation musicales, questionnaire de sensibilité à la récompense musicale, ainsi que l'Edinburgh Handedness Inventory. Les participants étaient ensuite installés dans une cage de Faraday afin de limiter le bruit électrique ambiant.

Les participants étaient assis à environ 50 cm de l'écran, face à un microphone cardioïde, équipés d'écouteurs intra-auriculaires pour la présentation des stimuli auditifs. Avant le début des tâches expérimentales, le montage EEG était réalisé à l'aide d'un système de 64 électrodes disposées selon le système international 10–20. Une fois l'installation terminée, Les consignes relatives aux différentes tâches étaient données oralement avant le début de l'enregistrement. La tâche d'amorçage rythmique était administrée en premier. Elle était suivie des tâches de tapping réalisées dans le même environnement expérimental. Ces tâches étaient présentées dans l'ordre suivant : tapping non cadencé, tapping synchronisé à un métronome, tapping synchronisé à des extraits musicaux (Fiveash et al., 2022), tapping sur structures binaires (duple) et ternaires (triple), puis BST (Beat Saliency Test ; Fujii & Schlaug, 2013) en perception et en production. Des pauses étaient proposées entre les différentes parties de la session et les consignes pour chaque épreuve étaient présentées sur l'écran avant chaque test.

La tâche principale reposait sur un paradigme d'amorçage rythmique non verbal. Chaque essai consistait à lire à voix haute une phrase immédiatement après l'écoute d'une amorce auditive. Le plan expérimental suivait une structure 2 × 2 × 2 entièrement intra-sujets, croisant la régularité de l'amorce (régulière vs irrégulière), sa structure (trois vs quatre accentuations) et la structure prosodique de la phrase (trois vs quatre groupes accentuels). Une session expérimentale comprenait huit blocs de 32 essais. Dans chaque bloc, une des quatre amorces rythmiques était associée à l'un des deux types de phrases. Chaque amorce était ainsi appariée une fois avec les phrases à trois groupes accentuels et une fois avec les phrases à quatre groupes accentuels, de sorte que chaque participant rencontrait deux fois chaque rythme au cours de l'expérience. La présentation des amorces était pseudo-randomisée entre participants afin d'éviter qu'une même amorce n'apparaisse dans deux blocs consécutifs. L'ordre des amorces et des listes de phrases était en outre contrebalancé entre participants, de façon à ce que chaque appariement amorce × structure de phrase apparaisse au moins une fois à chaque position au cours de l'expérience. Le programme de présentation des stimuli et d'enregistrement de la parole a été développé sous GNU Octave (version 5.2.0) à l'aide de la Psychophysics Toolbox Version 3 (PTB-3 version 3.0.17) et exécuté sous Ubuntu 20.04.3. Chaque essai débutait par une croix de fixation blanche sur fond noir,

affichée pendant la présentation de l'amorce. Immédiatement après la fin de celle-ci, la phrase cible apparaissait à l'écran et le participant la lisait à voix haute pendant que la production orale était enregistrée. L'essai se terminait automatiquement après 4,8 secondes. Une courte pause était proposée entre les blocs.

Après la tâche d'amorçage, les participants débutaient la seconde partie de l'étude mise en place pour évaluer leur compétence sensorimotrice. Pour les tâches de production motrice, ils tapaient avec l'index de la main dominante sur un dispositif de percussion, en utilisant uniquement les mouvements du doigt et en évitant autant que possible les mouvements du poignet. Il leur était demandé de garder le regard dirigé vers l'écran et de ne pas regarder leur main pendant l'exécution. La première tâche était un tapping non cadencé : les participants devaient produire un rythme spontané aussi régulier que possible pendant 60 secondes. Cette tâche était répétée deux fois. Venait ensuite un tapping synchronisé à un métronome. Les participants entendaient une séquence isochrone de 60 sons identiques, séparés par un intervalle inter-onset de 500 ms, et devaient taper en synchronie avec eux. Afin d'optimiser la précision de la synchronisation, il leur était suggéré d'attendre les premiers sons avant de commencer à taper. Cette tâche était également réalisée deux fois. Une troisième tâche consistait à synchroniser le tapping avec deux extraits musicaux présentant chacun un beat régulier de 500 ms. Les participants devaient taper sur le beat, défini comme la pulsation régulière de la musique, en maintenant des intervalles aussi constants que possible entre les frappes. Chaque extrait était présenté deux fois. Les participants réalisaient ensuite deux tâches de tapping avec accentuation, l'un binaire et l'autre ternaire. Les stimuli consistaient en séquences isochrones de sons percussifs dans lesquelles un son sur deux, ou un son sur trois, était accentué. Les participants devaient reproduire cette structure en modulant l'amplitude de leurs frappes en accord avec les accents perçus. Enfin, le BST issu de la H-BAT *(*Fujii & Schlaug, 2013) était administré dans sa version perception puis production. Dans la partie perceptive, les participants devaient identifier si la structure entendue était binaire ou ternaire à partir d'un paradigme adaptatif de type two-down one-up. Dans la partie productive, ils devaient reproduire la structure perçue par modulation de l'amplitude de leurs frappes. L'ordre des deux sous-parties était constant au sein de la session.

Traitement des données

Les productions orales recueillies dans la tâche d'amorçage rythmique ont fait l'objet d'une sélection, d'un traitement semi-automatique et d'une vérification manuelle suivant la même méthodologie que celle utilisée dans les chapitres 2 et 3 de la partie 2 de la thèse. Deux mesures ont été extraites. La première était le temps d'initiation de la parole, défini comme la durée séparant la fin de l'amorce auditive et le début de la production vocale. La seconde était le débit de parole, calculé en rapportant le nombre de syllabes à la durée totale de l'énoncé, pauses comprises lorsqu'elles étaient présentes. Ces deux mesures permettent de distinguer, d'une part, les effets de l'amorce sur la transition entre l'écoute du rythme et le déclenchement de la parole, et, d'autre part, les effets éventuels sur l'organisation temporelle de l'énoncé une fois la production engagée. Les mêmes critères d'exclusion que ceux utilisés dans l'étude précédente (partie 2, chapitres 2 et 3) ont été utilisés. Ainsi, sur 7680 enregistrements réalisés, 7651 ont été utilisés pour l'analyse du temps d'initiation (soit 99,62 %), 7064 ont été utilisés pour les analyses du débit de parole (soit 91,98 %).

En complément des données de parole, des mesures issues des tâches de tapping, du BST et des questionnaires musicaux ont été recueillies afin de documenter certaines différences interindividuelles dans les habiletés rythmiques, la pratique musicale et la sensibilité à la récompense musicale. La sensibilité à la récompense musicale a été évaluée à l'aide de la version française du Barcelona Music Reward Questionnaire (BMRQ ; Saliba et al., 2016). Ce questionnaire comprend 20 items, répartis en cinq sous-échelles de quatre items chacune : *Musical Seeking*, *Emotion Evocation*, *Mood Regulation*, *Sensory-Motor* et *Social Reward*. Les participants répondaient sur une échelle de Likert en cinq points, allant de 1 (« pas du tout d'accord ») à 5 (« tout à fait d'accord »). Les réponses ont ensuite été transformées, selon la procédure de cotation officielle de la version française, en scores T pour chacune des sous-échelles ainsi que pour un score global de récompense musicale (M = 50, ET = 10). Dans les analyses statistiques où ces scores étaient utilisés comme prédicteurs, ils ont été standardisés sous forme de z-scores, afin de faciliter l'interprétation et la comparaison des coefficients. Les informations relatives à la pratique et à la formation musicales ont été recueillies à l'aide d'un questionnaire spécifique administré en début de session. Ce

questionnaire permettait notamment d'obtenir un score de musicalité, ainsi que des variables plus ciblées relatives à la pratique musicale et à la durée de formation musicale.

Le traitement des données de tapping s'est déroulé en deux étapes : l'extraction des événements temporels d'intérêt à partir des stimuli et des signaux produits, puis le calcul des variables dérivées utilisées dans les analyses. Pour les tâches de tapping synchronisé à un métronome et de tapping synchronisé à une structure binaire ou ternaire, les temps des beats cibles ont été extraits par seuillage de l'enveloppe du signal audio de retour (*fed-back audio signal*). Cette enveloppe était obtenue en appliquant un filtrage passe-bas de Butterworth du quatrième ordre, à 200 Hz, en phase nulle, sur le signal rectifié. Pour les extraits musicaux, deux beats de référence (l'un proche du début, l'autre proche de la fin) ont été annotés manuellement à partir de la forme d'onde, en repérant les événements percussifs les plus saillants. Les autres beats ont ensuite été interpolés automatiquement à intervalle régulier à partir de ces deux points de repère. Les frappes produites par les participants ont été extraites à partir du signal brut provenant du transducteur piézoélectrique du dispositif de percussion. Dans ce signal, les frappes apparaissent sous la forme de bouffées oscillatoires de haute fréquence. Afin de supprimer la composante lente pouvant apparaître lorsque le doigt restait en contact avec la surface ou à la suite d'un rebond, le signal a d'abord été différencié. Les temps de frappe ont ensuite été détectés par seuillage de l'enveloppe du signal différencié, selon une procédure analogue à celle utilisée pour les beats du métronome. Une période réfractaire de 100 ms a été appliquée afin d'éviter la détection de rebonds comme frappes distinctes. Les résultats de la détection automatique ont enfin été vérifiés visuellement, et les détections erronées ont été retirées.

Pour la tâche de tapping non cadencé, la variable principale retenue était le coefficient de variation des intervalles inter-frappes, obtenu en divisant l'écart-type par la moyenne de ces intervalles. Cette mesure fournissait un indice de stabilité temporelle de la production spontanée. Pour les tâches de tapping cadencé, deux variables ont été extraites à partir de statistiques circulaires : la précision et la consistance de synchronisation (Dalla Bella et al., 2017b ; Fiveash et al., 2022). Les séries temporelles des frappes et des beats cibles ont été représentées sur une échelle polaire allant de 0 à 360°, où 0° correspondait au beat cible. Chaque frappe était convertie en angle en fonction de sa distance temporelle au beat le plus proche, rapportée à l'intervalle inter-stimulus. À partir de cette distribution angulaire, la

précision était calculée comme l'angle moyen (*relative phase*) entre les frappes et les beats cibles. Des valeurs négatives traduisent une anticipation du beat, tandis que des valeurs positives reflètent un retard. Afin de faciliter l'interprétation et la comparaison entre tâches, cette mesure angulaire a ensuite été convertie en secondes selon la formule suivante :

$$Accuracy_{time(s)} = Accuracy_{rad} * \frac{ISI}{2\pi}$$

où $Accuracy_{rad}$ correspond à la déviation angulaire moyenne exprimée en radians et ISI à l'intervalle inter-stimulus de la tâche considérée. La consistance correspondait à la longueur du vecteur résultant moyen, comprise entre 0 et 1 : une valeur proche de 0 traduit une forte dispersion des frappes autour du beat, tandis qu'une valeur proche de 1 traduit une synchronisation très stable. Les valeurs de consistance ont ensuite été soumises à une transformation logit afin d'être utilisées dans les analyses statistiques. Conformément aux recommandations de Dalla Bella et al. (2017b), les dix premières frappes de chaque essai ont été exclues de l'analyse, et un essai n'était conservé que s'il comportait au moins dix frappes restantes. La précision n'était calculée que lorsque l'hypothèse d'une distribution uniforme des angles pouvait être rejetée à l'aide d'un test de Rayleigh au seuil de .05. Sauf indication contraire, les valeurs de précision et de consistance ont ensuite été moyennées sur les deux répétitions de chaque tâche. Pour le BST, les seuils de perception et de production étaient obtenus directement à partir du paradigme adaptatif mis en œuvre pendant la tâche. Le seuil correspondait à la moyenne des six points de retournement de la procédure *two-down, one-up*. Les seuils ont été extraits à partir des fichiers de données enregistrés pendant l'expérience.

Les données EEG recueillies pendant cette session n'ont pas été intégrées aux analyses présentées dans ce chapitre. Elles ont été enregistrées afin de permettre, dans un second temps, l'étude du suivi neurophysiologique des structures rythmiques entendues avant la production. Les analyses envisagées porteront notamment sur la cohérence cerveau–stimulus aux fréquences correspondant aux régularités temporelles des amorces. L'objectif sera de déterminer si les amorces régulières induisent un suivi neuronal plus marqué que les amorces irrégulières, et si ce suivi est associé aux compétences de synchronisation sensorimotrice ou aux effets comportementaux observés sur l'initiation de la parole.

Analyses statistiques

L'ensemble des analyses a été conduit dans une logique visant à articuler deux niveaux : la parole produite après l'amorce et certaines différences interindividuelles relevant des habiletés rythmiques et du rapport à la musique.

Les effets de la régularité de l'amorce (irrégulière : IR vs régulière : R), de sa structure (3 accentuations : A3 vs 4 accentuations : A4) et de la structure prosodique de la phrase (3 groupes accentuels : S3 vs 4 groupes accentuels : S4) sur le temps d'initiation de la parole et sur le débit de parole ont été évalués au moyen de modèles mixtes bayésiens ajustés avec le package brms (Bürkner, 2017) sous R. Pour chaque variable dépendante, un modèle complet incluant l'interaction Régularité × Structure de l'amorce × Structure de la phrase a été estimé, de manière à tester non seulement un effet général de la régularité, mais aussi sa dépendance à la structure métrique de l'amorce et à la structure prosodique de la phrase cible. Un terme d'interaction entre la régularité et le score de musicalité, centré, a également été introduit, afin d'examiner si l'effet de la régularité variait selon certaines différences interindividuelles liées au profil musical. La structure hiérarchique des données était prise en compte au moyen d'effets aléatoires par sujet et par phrase, permettant de modéliser l'hétérogénéité interindividuelle et inter-items. La formule du modèle était donc celle-ci :

Variable dépendante ~ Régularité * Structure de l'amorce * Structure de la phrase + Régularité * Score de musicalité_c + (Régularité * Structure de l'amorce * Structure de la phrase |sb| Sujet) + (Régularité * Structure de l'amorce |st| Phrase)

Le temps d'initiation a été modélisé à l'aide d'une loi ex-gaussienne, ce choix permettant de prendre en compte l'asymétrie habituelle des distributions de latence. En complément de l'équation portant sur la moyenne, les paramètres de dispersion comportaient également un intercept aléatoire par sujet. Le débit de parole a été modélisé à l'aide d'une loi gaussienne. Des priors faiblement informatifs ont été utilisés afin de régulariser l'estimation sans imposer de contrainte forte sur les paramètres.

Comme dans les chapitres 2 et 3 de la partie 2, l'interprétation des effets reposait sur les distributions a posteriori, des moyennes marginales estimées au niveau de la population, ainsi que sur des contrastes calculés sur l'échelle de la réponse à partir des distributions a posteriori. Les résultats sont rapportés sous la forme d'estimations moyennes accompagnées de leurs intervalles de crédibilité à 95 %, ainsi que de probabilités directionnelles.

Les liens entre les deux variables de parole et les indices individuels de compétences musicales et de tapping ont été examinés au moyen d'analyses de corrélations bivariées bayésiennes robustes, en suivant l'approche proposée par Baez-Ortega (2018) et adaptée au package *brms* par Kurz (2019). Les variables incluses dans ces analyses comprenaient les indices de consistance et de précision aux tâches de tapping, le score global de musicalité, les seuils auditifs de la H-BAT en perception et en production ainsi que le contraste entre les conditions d'amorçage en ce qui concerne le temps d'initiation et le débit de parole. Il s'agit donc des contrastes R–IR global et R–IR par structure d'amorce, soit R–IR sous A4 et R–IR sous A3. Un tel contraste sera appelé « delta » dans la suite du manuscrit et reflète un indice de l'effet de l'amorçage sur les deux mesures de la parole étudiées.

Pour chaque couple de variables (ex. delta temps d'initialisation et seuil auditif de la H-BAT en perception), un modèle bayésien bivarié distinct a été ajusté. À chaque itération, seules les deux variables concernées étaient sélectionnées, puis standardisées sous forme de z-scores afin de placer l'ensemble des mesures sur une échelle commune. Chaque modèle incluait uniquement un intercept pour chacune des deux variables et estimait leur corrélation résiduelle à l'aide du paramètre rescor. Les observations présentant une donnée manquante pour l'une des deux variables du modèle n'étaient pas utilisées pour l'association correspondante, sans entraîner l'exclusion du participant des autres analyses corrélationnelles.

Le modèle principal utilisait une loi de Student pour la distribution des erreurs, afin de rendre l'estimation de la corrélation plus robuste aux valeurs atypiques et aux écarts à la normalité. Avant l'ajustement des modèles, des graphiques exploratoires de dispersion ont été produits pour chaque bloc de variables afin d'examiner visuellement la forme générale des associations et d'identifier d'éventuelles valeurs atypiques. Les modèles ont ensuite été ajustés avec *cmdstanr*, puis les distributions a posteriori du paramètre de corrélation ont été

extraites pour chaque association. La convergence des chaînes a été vérifiée à l'aide de l'indice Rhat. L'interprétation des résultats reposait sur la distribution a posteriori du coefficient de corrélation, résumée par son estimation moyenne, son intervalle de crédibilité à 95 %, ainsi que par la probabilité directionnelle (Pdir), définie comme la probabilité postérieure que l'effet soit positif ou négatif selon la direction dominante.

Résultats

Production de la parole

En ce qui concerne la production de la parole, deux mesures ont été analysées : le temps d'initiation et le débit. Les statistiques descriptives sont présentées dans le Tableau 1. Les distributions empiriques par condition sont présentées dans la Figure 1. Les principaux contrastes issus des modèles sont rapportés dans le texte.

Variable	Facteur	Niveau	Amorce régulière (moyenne et ET)	Amorce irrégulière (moyenne et ET)
Temps d'initiation	Structure de l'amorce	A3	762.2 (110.4) ms	803.0 (118.6) ms
		A4	766.9 (122.4) ms	770.1 (132.5) ms
Débit de parole	Structure de la phrase	S3	5.708 (0.619) syll/s	5.702 (0.558) syll/s
		S4	5.519 (0.540) syll/s	5.553 (0.572) syll/s

Tableau 1. Statistiques descriptives du temps d'initiation selon la structure de l'amorce et du débit de parole selon la structure prosodique de la phrase.

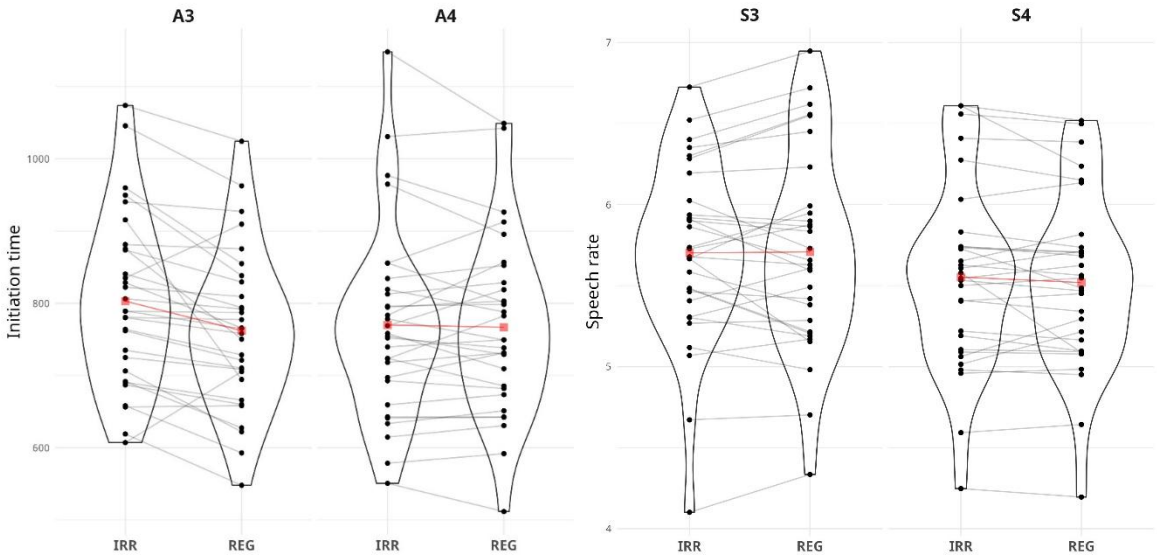

Figure 1. Résultats empiriques par condition selon la structure de l'amorce pour le temps d'initiation de la parole (ms, gauche) et selon la structure de la phrase pour le débit de parole (syll/s, droite). Les points noirs représentent les moyennes individuelles par condition ; les lignes grises relient les valeurs d'un même participant entre les conditions. Les diagrammes en violon représentent la distribution des observations. Les carrés rouges indiquent les moyennes de groupe, et les segments rouges relient ces moyennes entre les conditions. IRR : amorce irrégulière ; REG : amorce régulière ; A3 : Structure d'amorce à 3 accentuations ; A4 : structure d'amorce à 4 accentuations ; S3 : Structure de phrases à 3 groupes accentuels ; S4 : Structure de phrases à 4 groupes accentuels.

Temps d'initiation de la parole

Un effet global de la régularité a été observé sur le temps d'initiation de la parole : les participants commençaient à parler plus rapidement après une amorce régulière qu'après une amorce irrégulière, avec un contraste R–IR estimé à −20.66 ms, 95% CrI [−37.82, −4.09], Pdir = .99.

Cet effet moyen n'était toutefois pas homogène selon la structure de l'amorce. L'interaction entre la régularité et la structure de l'amorce était crédible, 33.54 ms, 95% CrI [14.70, 53.37], Pdir = 1.00, indiquant que l'avantage des amorces régulières dépendait du type de structure rythmique. Plus précisément, cet avantage était net pour les amorces A3, avec un contraste R–IR de −37.43 ms, 95% CrI [−58.39, −17.77], Pdir = 1.00. En revanche, aucun

effet clair de la régularité n'était observé pour les amorces A4, −3.88 ms, 95% CrI [−22.94, 14.66], Pdir = .66.

La structure de la phrase ne présentait pas d'effet global crédible sur le temps d'initiation. Elle modulait toutefois l'effet de la régularité dans certaines configurations. En particulier, pour les amorces A3, l'avantage de la régularité différait entre les phrases S3 et S4, 30.72 ms, 95% CrI [1.17, 60.63], Pdir = .98. Le bénéfice de la régularité était surtout marqué dans la configuration A3–S3 : les participants initiaient leur production plus rapidement après R qu'après IR, −52.79 ms, 95% CrI [−74.97, −30.87], Pdir = 1.00. Dans la configuration A3–S4, l'effet allait dans le même sens, mais il était plus incertain, −22.07 ms, 95% CrI [−50.76, 5.35], Pdir = .94. Enfin, aucun effet crédible de la musicalité n'a été observé sur le temps d'initiation.

Ainsi, l'effet global de régularité sur le temps d'initiation semble principalement porté par les amorces A3, et plus particulièrement par la configuration A3–S3.

Débit de parole

Concernant le débit, les phrases S4 étaient produites avec un débit plus lent que les phrases S3, avec un contraste S4–S3 estimé à −0.17 syll/s, 95% CrI [−0.26, −0.07], Pdir = 1.00. Cet effet était retrouvé dans l'ensemble des contrastes. En fonction de la structure de l'amorce, le contraste S4–S3 était crédible aussi bien pour les amorces A3 (−0.14 syll/s, 95% CrI [−0.25, −0.04], Pdir = .99), que pour les amorces A4, (−0.19 syll/s, 95% CrI [−0.30, −0.08], Pdir = 1.00). De même, en fonction de la régularité de l'amorce, les phrases S4 étaient produites plus lentement que les phrases S3 après amorces irrégulières, −0.15 syll/s, 95% CrI [−0.26, −0.03], Pdir = .99, comme après amorces régulières, −0.19 syll/s, 95% CrI [−0.29, −0.09], Pdir = 1.00.

Aucun effet moyen crédible de la régularité n'a été observé sur le débit de parole. Pour les phrases S3, une interaction entre régularité et structure de l'amorce a toutefois été observée, −0.10 syll/s, 95% CrI [−0.21, −0.00], Pdir = .98. Cette interaction indiquait que la différence entre amorces régulières et irrégulières n'était pas identique pour les structures A3 et A4. Les contrastes simples associés ne mettaient cependant pas en évidence d'effet crédible de la régularité au sein de chaque structure d'amorce.

Aucun effet crédible de la musicalité n'a été observé sur le débit de parole.

Corrélations entre les compétences rythmiques et les effets sur la production de la parole

La figure 2 représente les résultats observés dans les différentes tâches rythmiques réalisées par les participants. L'objectif de cette étude étant d'examiner l'association entre les capacités rythmiques et l'effet de l'amorçage rythmique sur la parole, nous ne développons pas davantage ces résultats.

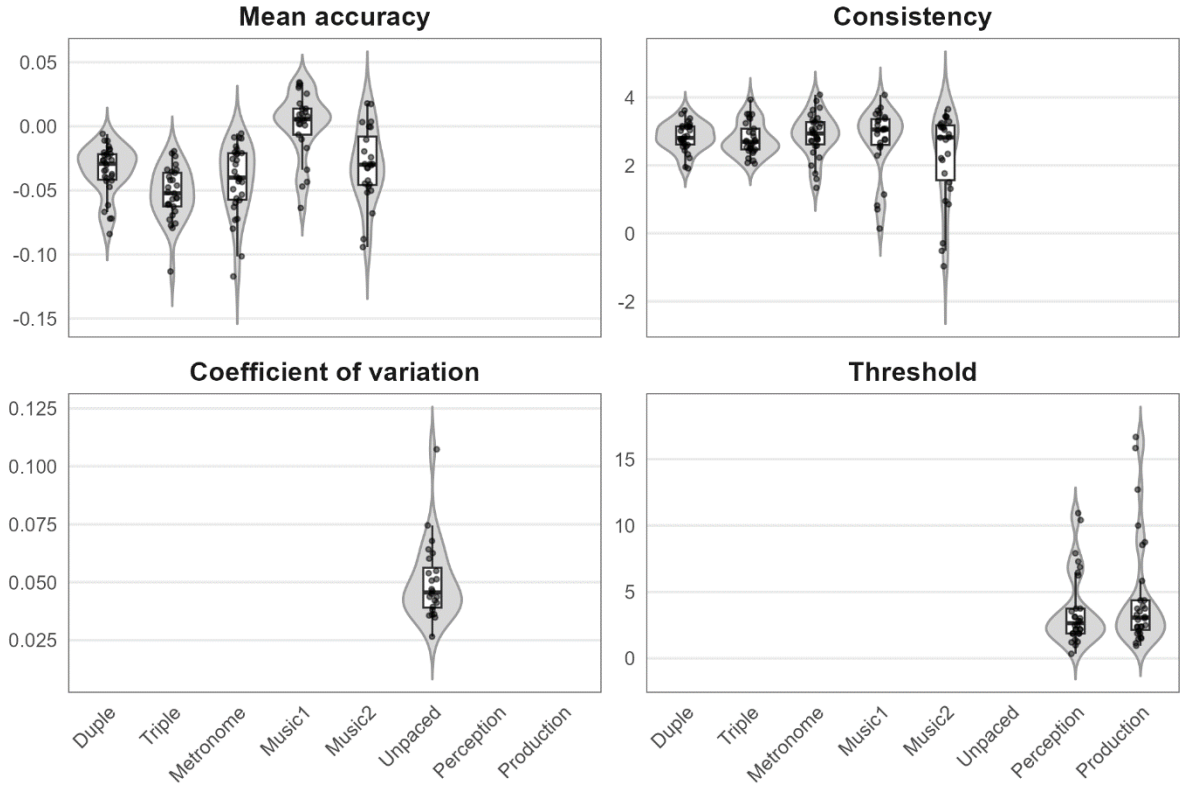

Figure 2 : Distribution des indices individuels de compétences rythmiques, musicales et auditives. Les graphiques en violon représentent la distribution des valeurs individuelles pour chaque tâche ou condition. Les points noirs correspondent aux valeurs individuelles des participants, et les boîtes représentent la médiane ainsi que l'intervalle interquartile. Les panneaux présentent respectivement la précision moyenne, la consistance, le coefficient de variation et les seuils de perception et de production de la structure métrique, issus du BST de la H-BAT. Les catégories représentées correspondent aux différentes tâches évaluées. Pour la tâche « unpaced », dans un souci de lisibilité, les valeurs extrêmes (3 valeurs supérieures à 0.2) ne sont pas illustrées sur la figure.

Compte tenu du nombre d'associations examinées et de la taille modérée de l'échantillon, les analyses corrélationnelles doivent être considérées comme exploratoires. Elles visaient à examiner les liens entre, d'une part, les mesures individuelles de synchronisation rythmique, le score global de musicalité lié à la formation et à la pratique musciales et les seuils auditifs de la H-BAT, et, d'autre part, les variables de parole. Les contrastes R–IR ont été calculés séparément pour les structures A3 et A4, afin d'évaluer si les différences individuelles étaient associées au bénéfice relatif des amorces régulières par rapport aux amorces irrégulières pour chaque structure d'amorce. Parmi l'ensemble des associations testées, une seule corrélation atteignait le critère de crédibilité. Elle concernait la consistance dans la tâche de tapping avec accentuation tous les deux temps et le delta du temps d'initiation pour les amorces de structure A4, avec une association négative faible à modérée, r = −0.36, 95% CrI [−0.65, −0.00], Pdir = .98. Ainsi, les participants présentant une synchronisation plus consistante tendaient à montrer un contraste R-IR plus négatif, c'est-à-dire un temps d'initiation plus court après l'amorce régulière que l'amorce irrégulière pour la structure A4.

Deux corrélations supplémentaires allaient dans un sens similaire, sans pour autant atteindre le critère de crédibilité. Elles concernaient également le delta du temps d'initiation pour la structure A4, avec la consistance dans la tâche triple, r = −0.35, 95% CrI [−0.64, 0.03], Pdir = .97, ainsi qu'avec la précision dans la tâche duple, r = −0.34, 95% CrI [−0.63, 0.03], Pdir = .97.

Dans l'ensemble, ces résultats ne mettent pas en évidence de relation généralisée entre les capacités rythmiques et l'impact de l'amorçage sur la parole. Ils suggèrent toutefois un signal convergent, mais limité, autour des liens entre certaines compétences rythmiques et la modulation du temps d'initiation pour les amorces de structure A4.

Discussion

Les résultats présentés dans ce chapitre visaient à préciser dans quelles conditions un rythme auditif non verbal présenté avant la parole peut influencer la production orale. Deux niveaux de production ont été distingués : le temps d'initiation, qui renseigne sur la transition entre l'écoute de l'amorce et le déclenchement de la parole, et le débit de parole, qui reflète davantage la dynamique temporelle globale de l'énoncé une fois la production engagée. Cette distinction permettait de déterminer si l'amorçage rythmique agit plutôt comme un contexte préparatoire facilitant l'entrée dans la parole, ou s'il modifie plus largement l'organisation temporelle de la phrase produite.

Dans l'ensemble, les résultats soutiennent davantage la première interprétation. Les amorces régulières étaient associées à des temps d'initiation plus courts que les amorces irrégulières, suggérant qu'une structure temporelle prédictible peut faciliter la préparation de la réponse vocale. Toutefois, cet effet n'était pas uniforme : il était crédible pour les amorces A3, mais pas pour les amorces A4, et apparaissait particulièrement net lorsque l'amorce A3 précédait une phrase S3. Le débit de parole suivait un profil différent : il était principalement déterminé par la structure prosodique de la phrase cible, les phrases S4 étant produites plus lentement que les phrases S3, indépendamment de la régularité ou de la structure de l'amorce. Enfin, les analyses corrélationnelles suggéraient un lien limité entre compétences rythmiques et effets comportementaux : une seule association crédible a été observée, entre la consistance dans la tâche de tapping avec accentuation tous les deux temps et le delta R–IR du temps d'initiation pour les amorces A4.

Un effet de régularité centré sur l'initiation de la parole

Le premier résultat important concerne le temps d'initiation. Les participants commençaient globalement à parler plus rapidement après une amorce régulière qu'après une amorce irrégulière. Ce résultat est cohérent avec celui de l'étude présentée dans le chapitre 2 de la partie 2 de la thèse, et compatible avec l'idée qu'un rythme régulier peut fournir un cadre temporel facilitateur, en rendant la séquence entendue plus prédictible et en favorisant la préparation de la réponse vocale. Cette interprétation rejoint la DAT, selon laquelle les régularités temporelles externes peuvent orienter l'attention vers des moments

attendus et améliorer le traitement d'événements ultérieurs (Large & Jones, 1999 ; Jones, 2019). Elle est également compatible avec le cadre PRISM, qui propose que les effets du rythme sur la parole et le langage reposent notamment sur l'extraction des régularités temporelles et sur leur articulation avec le système sensorimoteur (Fiveash et al., 2021).

Cependant, l'effet de la structure des amorces doit être pris en compte. La différence entre amorces régulières et irrégulières était crédible pour les amorces de structure A3, mais elle ne l'était pas pour les amorces de structure A4. À un niveau plus fin, cet effet était particulièrement net lorsque les amorces de structure A3 précédaient des phrases de structure S3. Lorsque ces mêmes amorces précédaient des phrases de structure S4, l'effet allait dans la même direction, mais restait plus incertain selon les critères de crédibilité retenus.

Cette observation est importante, car elle invite à distinguer la régularité comme propriété formelle du stimulus et la prédictibilité temporelle effectivement exploitable par le participant. Une amorce régulière peut fournir des repères temporels clairs, mais ces repères ne sont pas nécessairement utiles de la même manière selon la structure de l'amorce, la structure de la phrase cible et la relation entre les deux. Dans le cas présent, les essais dans lesquels une amorce régulière de structure A3 précédait une phrase de structure S3 semblent constituer une situation particulièrement favorable, probablement parce que la structure de l'amorce offrait des repères temporels compatibles avec l'organisation prosodique de la phrase à produire. Cette hypothèse devra toutefois être approfondie dans de futurs travaux, d'autant que le même profil n'a pas été observé lorsque les amorces régulières de structure A4 précédaient des phrases de structure S4.

L'absence de différence crédible entre amorces régulières et irrégulières pour les amorces de structure A4 doit également être interprétée avec prudence. Elle ne signifie pas nécessairement que la régularité ne joue aucun rôle pour cette structure. Elle peut aussi s'expliquer par le fait que l'amorce irrégulière correspondante ne constitue pas une condition défavorable au même titre que l'amorce irrégulière de structure A3. Si l'amorce irrégulière de structure A4 conserve certains repères temporels exploitables, ou si le positionnement plus

tardif de son dernier événement auditif facilite la préparation immédiate de la réponse, alors le contraste avec l'amorce régulière de même structure peut être atténué. Dans ce cas, l'absence de différence claire entre les amorces régulières et irrégulières de structure A4 ne remet pas en cause le rôle possible de la régularité, mais suggère que les conditions irrégulières ne sont pas toutes équivalentes.

Ainsi, les résultats soutiennent l'idée qu'une amorce régulière peut faciliter l'initiation de la parole, en particulier lorsque sa structure fournit des repères temporels suffisamment saillants et exploitables pour préparer la réponse vocale. Cette conclusion rejoint les travaux montrant que les effets de l'amorçage rythmique dépendent fortement des propriétés précises des stimuli, du type de tâche et de la relation temporelle entre amorce et cible (Cason & Schön, 2012 ; Cason et al., 2015 ; Fiveash et al., 2020 ; Przybylski et al., 2013).

Une absence d'effet généralisé sur le débit de parole

Le deuxième résultat important concerne le débit de parole. Contrairement au temps d'initiation, le débit n'était pas modulé de manière crédible par la régularité de l'amorce. Le résultat le plus robuste concernait la structure prosodique de la phrase : les phrases S4 étaient produites plus lentement que les phrases S3, et cet effet était retrouvé quelle que soit la régularité ou la structure de l'amorce. Cette différence pourrait refléter le coût temporel associé à une organisation en groupes accentuels plus nombreux, impliquant davantage de frontières prosodiques potentielles et de proéminences et pauses à réaliser. En dehors de cet effet de structure de phrase, le seul effet impliquant les propriétés de l'amorce concernait l'interaction entre régularité et structure d'amorce pour les phrases S3. Cette interaction indiquait que l'écart entre amorces régulières et irrégulières n'était pas identique pour les structures A3 et A4, mais les contrastes simples correspondants ne mettaient pas en évidence d'effet crédible de la régularité au sein de chaque structure d'amorce. Ce résultat doit donc être interprété avec prudence.

Cette dissociation entre temps d'initiation et débit suggère que l'amorçage rythmique agit davantage sur la préparation immédiate de la réponse que sur le déroulement temporel global de l'énoncé une fois la parole débutée. Autrement dit, l'amorce semble pouvoir faciliter l'initiation de la parole, sans pour autant modifier de manière stable le tempo de production.

Une fois la production commencée, le débit paraît surtout dépendre des propriétés propres de la phrase, notamment sa structure prosodique, la distribution des groupes accentuels et les contraintes de planification linguistique et articulatoire.

Ce profil rejoint, avec prudence, les résultats de l'étude précédente conduite auprès de populations cliniques, dans laquelle les effets de l'amorçage sur le débit étaient également limités et n'apparaissaient clairement que chez les personnes atteintes d'ASC. Dans le présent chapitre, mené auprès d'adultes neurotypiques, les données ne soutiennent donc pas l'existence d'un effet généralisé de l'amorce sur le débit de parole. Le résultat principal demeure que cette variable est davantage contrainte par la structure prosodique de la phrase cible que par les propriétés rythmiques de l'amorce.

Cette distinction peut être mise en relation avec les modèles de production de la parole, qui distinguent différentes étapes de préparation, de sélection séquentielle, d'initiation et de contrôle sensorimoteur de la production (Bohland et al., 2010 ; Tourville & Guenther, 2011 ; Guenther, 2016). Les résultats obtenus ici suggèrent que le rythme externe pourrait agir préférentiellement sur les processus préparatoires précédant l'entrée dans la parole, plutôt que sur l'exécution continue de l'énoncé. Cette interprétation reste toutefois spéculative, dans la mesure où les modèles actuels de production de la parole intègrent encore peu la manière dont des indices temporels externes, verbaux ou non verbaux, peuvent moduler l'état préparatoire du système avant la production.

Régularité, compatibilité temporelle et limites de l'interprétation en termes de congruence

L'un des apports du chapitre est de montrer que la régularité ne peut pas être interprétée indépendamment de la structure de l'amorce et de la phrase. L'hypothèse initiale pouvait laisser attendre un effet de congruence : une amorce structurée en trois accents faciliterait davantage une phrase S3, tandis qu'une amorce structurée en quatre accents faciliterait davantage une phrase S4. Les résultats ne confirment que partiellement cette logique.

La configuration A3–S3 présente effectivement le profil le plus compatible avec une hypothèse de correspondance structurelle : l'amorce régulière y est associée à des temps d'initiation plus courts que l'amorce irrégulière. Toutefois, l'absence d'effet comparable pour

A4–S4 empêche de conclure à un effet général de congruence. Il semble donc préférable de parler de compatibilité temporelle sélective plutôt que de congruence métrique au sens strict.

Cette nuance est importante. Les structures rythmiques non verbales et les structures prosodiques de la parole ne se correspondent pas nécessairement de manière directe. Une amorce peut partager avec la phrase un nombre d'accents ou une périodicité globale sans pour autant offrir des repères effectivement utiles pour la préparation de la production. À l'inverse, une amorce irrégulière peut contenir certains événements saillants ou certains repères temporels qui facilitent le passage à l'action. L'effet observé dépend donc probablement de plusieurs propriétés combinées : régularité globale, saillance des accents, position du dernier événement, stabilité locale du groupement et compatibilité avec la structure prosodique attendue.

Cela n'invalide pas l'effet de régularité observé pour A3, mais invite à éviter une interprétation binaire du type « régulier = facilitateur » et « irrégulier = défavorable ». Les résultats suggèrent plutôt que l'effet d'une amorce dépend de la configuration temporelle complète du stimulus et de la manière dont cette configuration peut être utilisée pour préparer la réponse vocale. De futurs travaux gagneraient à manipuler séparément la régularité, la structure métrique, la position du dernier événement et le délai entre la fin de l'amorce et l'apparition de la phrase.

Rôle des compétences rythmiques et musicales

Un autre objectif de cette étude était d'examiner si les différences individuelles rythmiques ou musicales pouvaient être associées à la sensibilité comportementale à l'amorçage. Dans la logique du cadre PRISM, les effets du rythme sur la parole ne dépendent pas seulement des propriétés du stimulus, mais aussi de la capacité du participant à extraire une régularité temporelle et à la convertir en ressource pour l'action (Fiveash et al., 2021). Les tâches de tapping constituaient ici une manière d'approcher ce versant sensorimoteur.

Les résultats corrélationnels restent toutefois limités. Une seule association crédible a été observée : la consistance dans la tâche de tapping avec accentuation tous les deux temps était associée au delta R–IR pour la structure A4 du temps d'initiation. Cette association négative, faible à modérée, suggère que les participants les plus consistants dans cette tâche tendaient à montrer un bénéfice plus marqué de R par rapport à IR. Ce résultat est compatible

avec l'idée que de meilleures capacités de synchronisation peuvent favoriser l'exploitation d'une structure rythmique pour préparer la parole. Il rejoint plus largement les travaux montrant que les compétences de perception et de production du beat varient entre individus et peuvent présenter des profils dissociés selon les dimensions évaluées (Dalla Bella et al., 2017b ; Fujii & Schlaug, 2013). Il est également cohérent avec l'idée que la synchronisation auditivo-motrice s'inscrit dans un ensemble plus large de capacités de traitement temporel et de prédiction rythmique (Nozaradan et al., 2016).

Cependant, cette association demeure isolée. Aucune relation crédible n'a été observée entre le score global de musicalité et les variables de parole, ni entre les autres indices rythmiques et les effets comportementaux. Les résultats suggèrent donc une relation possible, mais fragile et dépendante de la mesure considérée. La consistance dans une tâche de tapping accentué pourrait refléter une compétence particulièrement pertinente pour la tâche, parce qu'elle implique non seulement une synchronisation régulière, mais aussi la prise en compte d'une structure accentuelle. Cette dimension se rapproche davantage de la structure des amorces utilisées dans l'expérience qu'un simple maintien de tempo ou qu'un score global de musicalité.

Ces résultats doivent toutefois être replacés dans une littérature montrant que les compétences rythmiques ne constituent pas un construit unitaire, mais regroupent plusieurs composantes partiellement dissociables. Les travaux de Fiveash et al. (2022), fondés sur plusieurs tâches de perception et de production rythmiques, distinguent notamment des dimensions liées à la production rythmique, à la perception du beat et à la mémoire de séquences rythmiques. Dans cette perspective, la capacité à synchroniser une action avec une structure externe, la capacité à percevoir une organisation métrique, la pratique musicale et la capacité à mobiliser ces informations dans une tâche de parole ne se recouvrent probablement pas entièrement. Cette dissociation pourrait contribuer à expliquer pourquoi les effets d'amorçage rythmique varient entre individus et restent parfois difficiles à prédire à partir d'un seul score global.

Limites et perspectives

Plusieurs limites doivent être soulignées. La première concerne le périmètre des variables comportementales analysées. Le temps d'initiation et le débit de parole permettent

de distinguer la mise en route de la réponse et le tempo global de l'énoncé, mais ils ne couvrent pas toute la complexité de la prosodie produite. Des effets plus fins pourraient apparaître sur la distribution des pauses, la régularité des intervalles syllabiques, la réalisation des proéminences ou les variations mélodiques, qui n'ont pas été analysées dans cette étude.

La deuxième limite concerne la définition des amorces irrégulières. Les résultats montrent que, selon leur structure (A3 ou A4), elles n'induisent pas le même effet. Cela limite l'interprétation du facteur « régularité » comme opposition simple entre stimuli réguliers et irréguliers. De futurs travaux devraient caractériser les amorces à partir d'indices continus de régularité locale, de saillance du beat, de stabilité du groupement ou d'imprévisibilité temporelle, afin de mieux décrire ce que chaque amorce rend effectivement disponible pour le système perceptivo-moteur.

Enfin, les analyses corrélationnelles restent exploratoires. Le nombre d'associations examinées est important au regard de la taille de l'échantillon, et seule une association entre compétences rythmiques et parole répondait aux critères retenus. Les résultats doivent donc être considérés comme des indices plutôt que comme des conclusions définitives. Ils suggèrent néanmoins que les compétences de synchronisation, en particulier lorsqu'elles impliquent une structure accentuelle, pourraient contribuer à expliquer une partie de la variabilité interindividuelle dans les effets de l'amorçage rythmique.

Conclusion

Les résultats obtenus montrent que l'amorçage rythmique non verbal peut influencer la préparation de la parole au niveau du temps d'initiation. Les amorces régulières sont associées à une initiation plus rapide que les amorces irrégulières, mais cet effet est sélectif en fonction de la structure de l'amorce. Le débit de parole, en revanche, apparaît surtout contraint par la structure prosodique de la phrase cible.

Ces résultats invitent à dépasser une lecture simple de la régularité. L'effet d'une amorce rythmique dépend de la structure temporelle disponible, de sa compatibilité avec la phrase à produire et du positionnement des événements auditifs. Les analyses

corrélationnelles suggèrent par ailleurs que les compétences de synchronisation rythmique peuvent être liées à certains effets comportementaux, mais ce lien reste limité et exploratoire.

Ce chapitre contribue ainsi à préciser les conditions dans lesquelles un rythme perçu peut devenir une ressource pour la parole produite. Il montre que le rythme n'agit pas comme un facilitateur uniforme, mais comme un contexte temporel dont l'efficacité dépend de la manière dont le système perceptivo-moteur peut l'exploiter pour préparer l'action vocale. Au-delà de ces apports expérimentaux, ces résultats soulignent l'importance d'une caractérisation fine des stimuli rythmiques et des mécanismes qu'ils sollicitent, notamment dans la perspective d'une utilisation raisonnée de ce type d'approche en contexte clinique.

Chapitre 5
De la théorie neurocognitive à la pratique orthophonique

Les chapitres précédents ont permis d'examiner, sous des angles complémentaires, la manière dont le rythme peut interagir avec la production de la parole. Le présent chapitre prolonge cette réflexion en déplaçant le regard vers la pratique clinique. L'objectif est d'examiner les conditions dans lesquelles ces approches sont effectivement connues, utilisées et appropriées par les professionnels de terrain.

L'article présenté dans ce chapitre, intitulé *From neurocognitive frameworks to clinical practice: how music and rhythm are utilized by speech-language pathologists* et publié dans *American Journal of Speech-Language Pathology* en 2026, s'inscrit dans cette perspective. Fondé sur une enquête par questionnaire auprès d'orthophonistes exerçant en français, il vise d'une part à dresser un état des lieux des usages cliniques de la musique et du rythme, en particulier de leurs dimensions non verbales, et d'autre part à identifier les facteurs susceptibles d'être associés à leur emploi.

Les résultats montrent que la musique et le rythme sont effectivement utilisés dans la pratique orthophonique pour une diversité de troubles, avec une présence particulièrement marquée dans les troubles développementaux du langage, mais aussi dans l'aphasie et la dysarthrie. Ils mettent également en évidence un contraste important entre, d'un côté, l'intérêt manifesté pour ces approches et leur usage relativement fréquent, et, de l'autre, le sentiment largement partagé de manquer de formation, d'outils stabilisés et d'appuis suffisamment solides pour les mettre en œuvre dans un cadre véritablement fondé sur les preuves. L'étude souligne en outre que leur utilisation est associée à plusieurs facteurs convergents, notamment la formation clinique spécifique, la familiarité avec la littérature scientifique et l'engagement actuel dans des activités musicales ou rythmiques, tandis que la formation musicale extrascolaire antérieure ne semble pas constituer, à elle seule, un facteur déterminant.

Le résumé en français et l'article dans sa version publiée sont présentés ci-après.

Brieussel, J., Fumel, J., Ott, L., & Basirat, A. (2026). From Neurocognitive Frameworks to Clinical Practice: How Music and Rhythm Are Utilized by Speech-Language Pathologists. *American journal of speech-language pathology*, 1–13. Advance online publication. https://doi.org /10.1044/2025_AJSLP-25-00234

Résumé

Objectif : les interventions fondées sur la musique et le rythme dans le domaine de l'orthophonie ne sont pas nouvelles, et on observe une augmentation des études neuroscientifiques visant à expliquer leurs effets sur le plan mécanistique. Cependant, la manière dont les orthophonistes intègrent ces recherches en constante évolution dans leur pratique clinique quotidienne reste difficile à cerner. Cette étude vise à dresser un état des lieux et à identifier les facteurs susceptibles d'être associés à l'utilisation de ces méthodes.

Méthode : les données ont été recueillies auprès d'orthophonistes de pays francophones via un questionnaire en ligne. Elles ont été analysées à l'aide de méthodes statistiques descriptives et associatives.

Résultats : cent trente-cinq réponses valides ont été analysées. Bien qu'ils ne disposent pas d'une formation clinique suffisante aux méthodes basées sur la musique et le rythme, les orthophonistes ont déclaré utiliser la musique et le rythme pour traiter divers troubles. Dans ce contexte, l'objectif principal était l'amélioration spécifique des symptômes, et le choix de la méthode dépendait largement du diagnostic clinique du patient. L'amélioration de la formation clinique, le renforcement des connaissances fondées sur des preuves grâce à la consultation de la littérature scientifique et la promotion d'un engagement personnel envers la musique se sont avérés être des facteurs cruciaux pour améliorer l'intégration du rythme et de la musique dans la pratique clinique.

Conclusion : les études futures devraient s'attacher à identifier des moyens plus efficaces de traduire les connaissances théoriques issues de la recherche scientifique en outils cliniques concrets pour les interventions fondées sur la musique et le rythme. De plus, il convient de combler certaines lacunes importantes dans la recherche, notamment en évaluant l'efficacité de ces interventions, en comprenant les mécanismes sous-jacents à leurs bienfaits potentiels et en examinant comment les orthophonistes évaluent leur efficacité dans leur pratique quotidienne.

Mots-clés : rythme et musique ; intervention ; orthophonie

From neurocognitive frameworks to clinical practice: how music and rhythm are utilized by speech-language pathologists

Johanna Brieussel[1], Jules Fumel[2], Laurent Ott [2] and Anahita Basirat[2]

1. Univ. Lille, Centre de Formation Universitaire en Orthophonie, Département de Médecine, UFR3S, F-59000 Lille, France

2. Univ. Lille, CNRS, UMR 9193 - SCALab - Sciences Cognitives et Sciences Affectives, F-59000 Lille, France

Abstract:

Purpose: Music- and rhythm-based interventions in the field of speech and language pathology are not new, and there has been an increase in neuroscientific studies seeking mechanistic accounts of their effects. However, the way speech-language pathologists (SLPs) are integrating the evolving research into real-world clinical practice remains elusive. This study aims to provide an overview of the situation and to identify factors that might be associated with the use of these methods.

Method: Data were collected from SLPs in French-speaking countries via an online survey. They were analyzed using descriptive and associative statistical methods.

Results: One hundred and thirty-five valid responses were analyzed. Despite not having enough clinical training in music- and rhythm-based methods, SLPs reported using music and rhythm to treat various disorders. When doing so, the primary goal was symptom-specific improvement, and the choice of method was largely dependent on the patient's clinical diagnosis. Improving clinical training, enhancing evidence-based knowledge through engagement with the scientific literature, and fostering personal engagement with music were found to be crucial factors in improving the integration of rhythm and music into clinical practice.

Conclusion: Future studies should focus on identifying more effective ways to translate theoretical insights from scientific research into practical clinical tools for music- and rhythm-based interventions. Moreover, critical gaps in research must be addressed, including evaluating the effectiveness of these interventions, understanding the mechanisms underlying their potential benefits, and exploring how SLPs assess their efficiency in everyday practice.

Keywords: Rhythm and music; intervention; speech and language pathology

1. Introduction

For decades, music- and rhythm-based approaches have been explored as tools to enhance therapeutic outcomes in rehabilitation science (for reviews, see Sihvonen et al., 2017; Cancer & Antonietti, 2022; Van Tellingen et al., 2023). Alongside these applied explorations, theoretical frameworks in the fields of cognitive psychology and cognitive neuroscience have evolved. These frameworks refine our understanding of the neurocognitive interplay between music and language and suggest specific mechanisms underlying the therapeutic benefits of such interventions. Prominent frameworks include OPERA (Overlap, Precision, Emotion, Repetition and Attention, Patel, 2011, 2014), the SEP hypothesis (Sound Envelope Processing and Synchronization and Entrainment to a Pulse hypothesis; Fujii & Wan, 2014) and PRISM (Processing Rhythm in Speech and Music, Fiveash et al., 2021). According to the OPERA framework, rehabilitation using instrumental music can benefit speech processing through brain plasticity, since music and speech share common cognitive and sensory processing mechanisms, mediated by overlapping brain networks, with music mobilizing precise auditory processing, focused attention, repetition, and emotional features. The SEP hypothesis emphasizes the different neural circuits involved in perceiving and producing speech rhythm and highlights the role of rhythm in rehabilitating disorders such as aphasia, stuttering, and Parkinson's disease. Partly based on these two theoretical frameworks, the PRISM model proposes three mechanisms that music and speech rhythm processing share: (a) the ability to perceive fine-grained variations in rhythm, pitch or timbre; (b) synchronization of neuronal oscillations to external rhythmic stimuli, known as neural entrainment; and (c) sensorimotor coupling, which involves the connection between auditory and motor cortices. According to the PRISM model, these mechanisms can be trained through music, which may lead to improvements in speech and language processing for disorders such as stuttering, dyslexia, and developmental language disorder. While these neurocognitive frameworks do not address rehabilitation methods specifically, they provide a foundation for understanding how music- and rhythm-based interventions could be useful in clinical settings.

Despite theoretical advancements, how these insights translate into real-world clinical practice remains elusive. Key questions remain regarding the perception and implementation of methods that utilize music and rhythmic elements by clinicians. Additionally, questions persist concerning the facilitators and barriers that influence the development and use of

these methods. One way to address this issue is to examine research on therapeutic interventions that incorporate music and rhythmic elements and the involvement of speech-language pathologists (SLPs) in these interventions (Ong et al., 2024). This approach, based on systematic or scoping literature reviews, provides valuable insight into SLPs' contributions to clinical research related to music and rhythm. However, it may not fully capture how such methods are applied in real-world clinical practice outside of controlled studies. The present study used a survey-based methodology to address this issue. SLPs are uniquely positioned to bridge the gap between theoretical insights and their practical implementation, and their input may help identify potential challenges in these interventions.

The abovementioned theoretical frameworks and experimental studies linking music and language focused on various pathologies and disorders such as aphasia (e.g., Stahl et al., 2011), stuttering (e.g., Frankford et al., 2021), developmental language disorder (e.g., Ladányi & al., 2021), dyslexia (e.g., Colling et al., 2017), hearing impairment (e.g., Cason et al., 2015), and Parkinson's disease (e.g., Kotz & Gunter, 2015). Such studies typically aim to investigate the underlying neurocognitive mechanisms by exploring how music and rhythm influence speech and language. For instance, Stahl et al. (2011) examined the roles of melody and rhythm in speech production among individuals with non-fluent aphasia. They compared various conditions including singing and rhythmic speech, and found that rhythm rather than melody, is essential for speech production, especially in individuals with basal ganglia lesions. The authors suggested that the benefits often attributed to melodic intoning in singing may stem from rhythm. They also emphasized the importance of lyric memory and motor automaticity in facilitating speech production.

Several studies have examined the effectiveness of music- and rhythm-related speech-language pathology intervention methods. Melodic Intonation Therapy (MIT) (Sparks & Holland, 1976) is one of the most widely used methods for treating acquired neurogenic communication impairments in adults (Ong et al., 2024). However, the absence of standardized protocols, the divergent outcomes, and inconsistent reporting may have diminished the reliability of the evidence supporting this method (Ree & al., 2025). Other methods, such as the pacing board (Van Nuffelen & al., 2010) and Rhythmic Speech Cueing® (Mainka & Mallien, 2014), have also shown promising results. Finally, the benefits of joint

interventions by music therapists and SLPs are also beginning to be explored (Johnson et al., 2019; Ong et al., 2024).

Little research has examined how SLPs integrate music- and rhythm-based interventions into their everyday clinical practice. Behaghel and Zumbansen (2022) did a survey to examine the use of melodic therapies and singing-based interventions for acquired neurogenic communication disorders. Although SLPs widely recognized the value of melodic therapies, their implementation was limited by insufficient training and inconsistent adherence to structured protocols. Moreover, the results showed that most respondents were aware of the therapeutic potential of singing, yet group singing was rarely used. In another survey-based study, Daigmorte et al. (2022) investigated SLPs' conceptualization and use of rhythm. Although they reported frequently using rhythm-based interventions to address oral language disorders and intuitively recognized rhythm as foundational to language development, they seldom explicitly associated it with linguistic function. Many also reported having a limited understanding of the theoretical principles and practical applications of rhythm in clinical contexts. While these studies offer valuable insights, they lack a comprehensive overview of the use of music- and rhythm-based interventions in SLP practice and the facilitators and barriers to their implementation in everyday practice.

In summary, a growing body of literature is exploring how music and rhythm may beneficially influence speech and language processing through the study of neurocognitive mechanisms. Although the use of music- and rhythm-based interventions in rehabilitation is not new, there has also been an increase in studies investigating the scientific basis of these methods (e.g., Ree & al., 2025; Cancer & Antonietti, 2022). However, it remains unclear how this evolving research, along with its associated questions and challenges, is being integrated into the clinical practice of SLPs. This study aimed to address this gap by investigating the use of music- and rhythm-based interventions in SLPs' clinical practice using a survey-based methodology. We were particularly interested in the use of non-verbal features related to rhythm and music. Therefore, participants were asked not to consider singing-based techniques when answering the questionnaire. The primary objective was to provide a comprehensive overview of current clinical practices involving music- and rhythm-based interventions, and SLPs' opinions of such methods from both clinical and scientific perspectives. The second objective was to identify factors associated with the use of these methods. Specifically, we explored the role of two

categories of factors: (a) personal factors unrelated to clinical training (e.g., formal extracurricular training and engagement in music or rhythmic activities), and (b) professional factors (e.g., clinical training in music- and rhythm-based interventions and familiarity with the scientific literature on such approaches in rehabilitation).

2. Methods

2.1. Participants

The target population was SLPs who practice in the French language, regardless of their country of certification and practice. Eligibility was determined through self-reporting using two questions at the beginning of the survey. If the responses failed to meet the criteria, the survey was automatically terminated. Participants who completed the entire questionnaire were included in the study (n = 135). No incentives were provided for participation. The study was conducted via an online survey, with responses collected anonymously (see Section 2.3). Completion of the questionnaire implied participants' informed consent. Approval was obtained from the University of Lille Data Protection Officer, who granted an exemption (N°1117).

2.2. Survey design

The questionnaire consisted of 38 questions divided into five sections and designed to obtain voluntary, anonymous responses. The sections were organized as follows: (1) SLPs' academic, professional and musical background; (2) profile of SLP's patients; (3) SLP's training and theoretical knowledge related to music- and rhythm-based methods; (4) use of music and rhythm in practice, including potential facilitators and barriers to using these methods; and (5) contexts in which these methods are used. Only participants who answered that they use music and rhythm in their practice answered the last two sections, except for the question regarding barriers, to which all participants were invited to respond. An information note appeared before the third section to define the terms 'music', 'simple auditory rhythm' and 'rhythmic movements' used throughout the survey. The term 'music' was defined as instrumental, thereby excluding singing; 'simple auditory rhythm' referred to rhythms such as those produced by a metronome; and 'rhythmic movements' encompassed gestures like tapping or more extensive movements. Regarding the disorders included in the survey, we

followed the OPERA (Patel, 2011, 2014), PRISM (Fiveash et al, 2021), and SEP (Fujii & Wan, 2014) frameworks, reviewed in the Introduction, and the musical sonification framework for handwriting (Véron-Delor et al., 2020). The disorders included in the survey were aphasia, developmental language disorder, dysarthria, dysgraphia, dyslexia, hearing impairment, and stuttering. The French terms used in the questionnaire were as follows: A*phasie*, *trouble développemental du langage*, *dysarthrie*, *dysgraphie*, *dyslexie*, *surdité*, and *bégaiement*. SLPs were also allowed to provide open responses if they use music- and rhythm-based methods for other disorders.

The design of the questionnaire was in line with the criteria of the Survey Checklist Manifesto regarding the formulation, formatting and organization of items (Gehlbach & Artino, 2018). There were 19 mandatory questions and 14 conditional ones. They included 16 single-choice questions, 11 multiple-choice questions, three Likert scales, and eight open-ended questions. Respondents could go back and modify their answers. One SLP, a music professional, and three final-year SLP students supervised the first draft of the questionnaire and provided feedback for subsequent revisions. An English version of the survey is available in Supplementary Material 1.

2.3. Data collection

The questionnaire was administered anonymously using LimeSurvey (http://www.limesurvey.org) via a website hosted on the servers of the University of Lille. The survey link was distributed via email to SLP unions and associations, social media groups dedicated to speech-language pathology, and personal and professional networks. Data collection took place between January and April 2025. The answer rate could not be calculated since the number of recipients to whom the questionnaire link was sent to was unknown.

2.4. Data analyses

The data were exported from LimeSurvey and analyzed using R software. Responses from participants who did not meet the inclusion criteria or who did not complete the survey to the end were excluded. First, we conducted descriptive analyses of the responses to generate an overview. For multiple-choice questions, we calculated the number and percentage of

responses to each option. For open-ended items, the data were analyzed using the conventional content analysis approach (Hsieh & Shannon, 2005). Each response was categorized, and the number and percentage of responses in each category were calculated. Classifications of disorders (questions 17, 18 and 28) were proposed by the first author (J.B) and reviewed by the second one (J.F), who is an SLP. Second, we qualitatively analyzed the open descriptions provided by participants regarding the methods they used (question 29). This analysis aimed to identify the most commonly employed approaches and the aspects that SLPs consider most important in their practice. Third, to determine factors potentially associated with the use of music- and rhythm-related methods, we analyzed the associations between their use (a 'yes-no' variable) and three factors via chi-square tests: (1) extracurricular music- and rhythm-related training (four 'yes-no' questions), (2) current music- and rhythm-related practices (three 'yes-no' questions), and (3) clinical initial training or continuing education on SLP methods involving music and rhythm (three 'yes-no' questions). For each factor, a 'yes' response to any constituent question was coded as 'yes' for that factor. Additionally, we evaluated the association between the use of these methods and familiarity with the scientific literature on music- and rhythm-based approaches in rehabilitation using a Kendall rank correlation. Responses to the scientific familiarity question (Likert-type) and use of music- and rhythm-based methods ('yes-no' question) were treated as ordinal variables in this analysis. Statistical significance was set at $p < .05$.

3. Results

We collected 194 responses, of which 135 met the inclusion criteria. Participants who did not provide French-language assessment or treatment, or did not currently practice as SLPs (n=6), were excluded, as were those who did not complete the survey (n=53).

3.1. Characteristics of participants and their patients

This section describes the demographic characteristics of the respondents. It also includes the patients' characteristics and information about the SLPs' personal musical and rhythmic training and practices. A summary of the demographic characteristics is presented in Table 1. The SLPs in the study were certified between 1978 and 2024, and had graduated from seven

different countries. One respondent graduated from two countries, and one answer was not interpretable. Most had studied in France (63%) or Belgium (21%). They practiced in seven different countries, primarily in France (81%). Most respondents had a private practice. Regarding the patients' profiles, most SLPs treated children and adolescents. Most of the participants were more often involved in treatment than assessment. One participant did not respond to the question about the allocation of time between treatment and assessment in their practice.

Table 1. Characteristics of participants. N and % represent number and percentage of participants.

		N	%
Year of certification			
	1978-1987	10	7
	1988-1997	18	13
	1998-2007	31	23
	2008-2017	34	25
	2018-2024	42	31
Mode of practice			
	Private practice	93	69
	Salaried	33	24
	Mixed practice	9	7
Type of practice			
	More treatment	113	84
	More assessment	19	14
	Balanced distribution	2	1
	Unanswered	1	1
Patient age profile			
	Mostly children and adolescents	93	69
	Mostly adults	24	18
	Balanced distribution	18	13

Participants were asked to estimate their experience with seven disorders, based on the aforementioned theoretical frameworks. They reported the highest level of experience with dyslexia and developmental language disorder, followed by aphasia and dysarthria (Figure 1). Using open-ended questions, we also asked which disorders they had most frequently evaluated and treated over the past twelve months. The results are presented in Supplementary Material 2. In summary, the most frequent assessments were in the following categories: written language disorders, spoken language disorders and neurological disorders

(30%, 28% and 13% of the answers, respectively). The most frequent treatments were in the same categories as the assessments (30%, 28% and 14%, respectively).

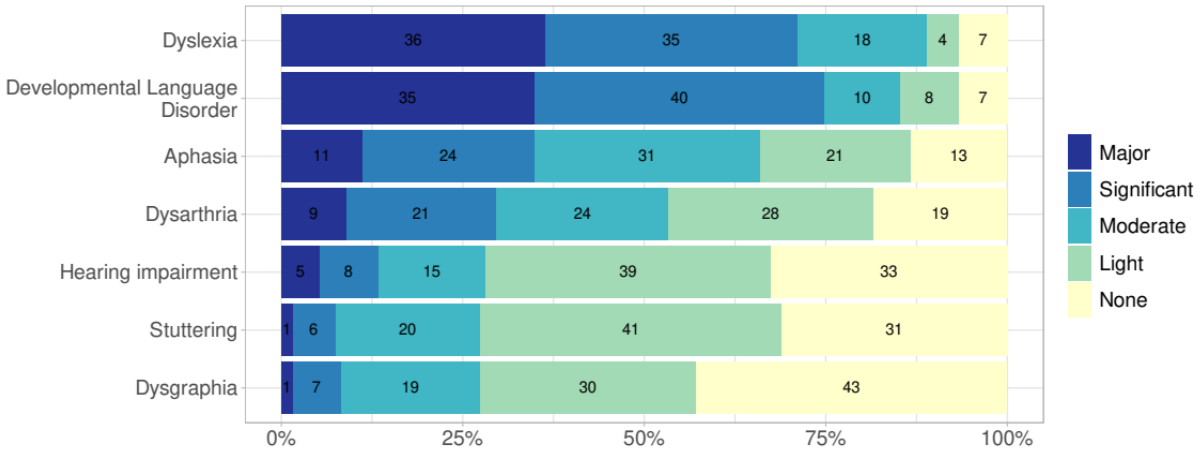

Figure 1. Clinical experience of participants in disorders for which theoretical frameworks report a beneficial effect of music and rhythm. For each disorder, numbers indicate the proportion of participants who selected each response category, i.e., major, significant, moderate, light, or none.

Regarding their previous extracurricular musical and rhythmic practice, 62% had studied an instrument, 61% had training in music theory or *solfège*, 49% had taken dance classes and 39% had received singing training. Overall, 83% of participants had benefited from at least one of these programs. Regarding their current music- or rhythm-based activities, 56% played an instrument, danced or sang. The most frequently reported activities were playing an instrument only (21%), followed by singing only (12%), dancing only (9%), and a combination of instrument playing and singing (9%). Next, we examined SLPs' self-reported musical and rhythmic skills. Ninety-four percent of participants reported that they could clap their hands in time to music, and 89% reported that they could dance in time to music.

3.2. Status of music- and rhythm-based approaches in SLP training and clinical practice

This section presents responses to questions about clinical aspects related to music and rhythm. Fifty-five percent of SLPs reported having benefited from training in the use of music, simple auditory rhythm, or rhythmic movement, either during their initial training or through post-certification continuing education. The training they most received was for rhythmic movements (37%) and simple auditory rhythm (36%), followed by music (29%). Of the SLPs who received training in at least one of these areas (n = 75), 20% answered 'often', 45% 'sometimes', 22% 'rarely' and 12% 'never' when asked about the frequency of the use of this training in their practice. One person did not answer the question. Regarding familiarity with research on rehabilitation methods based on music, simple auditory rhythm, and rhythmic movement, 38% of participants felt not familiar at all, 39% unfamiliar, and 19% moderately familiar. Only 4% felt quite familiar with these studies, and 1% felt very familiar.

Regarding clinical practice, SLPs were asked whether they used or had used music, simple auditory rhythm, or rhythmic movement in their work. Most of them confirmed using these types of interventions (70%). These participants (n=95) were then invited to answer a series of questions regarding their practice. First, they were asked for which disorder they were using or had used music- and rhythm-based methods. The most common disorders for which these methods were used were developmental language disorder (65%), followed by aphasia (45%) and dysarthria (43%) (Figure 2). They had the possibility to add other disorders. Thirty SLPs added a total of 39 responses. The most frequent response was in the voice disorders category (see Supplementary Material 2). SLPs who reported using music and rhythm in their practice were asked to briefly describe their method(s). Seventy-five participants responded to this question. The most representative responses were as follows: tapping and clapping were frequently mentioned techniques that were primarily used to cue syllables or letters. Another frequently reported technique was *Thérapie Mélodique et Rythmée* (TMR), an adaptation of the MIT for French-speaking patients (Eeckhout & Allichon, 1978). SLPs also reported using vocal techniques including singing, scat singing, and beatboxing, as well as nursery rhymes and songs, during their sessions. The use of musical instruments and metronomes was also mentioned, although less frequently than the methods listed above. Interestingly, many SLPs described incorporating methods that involve bodily movement, such as dancing or walking in rhythm.

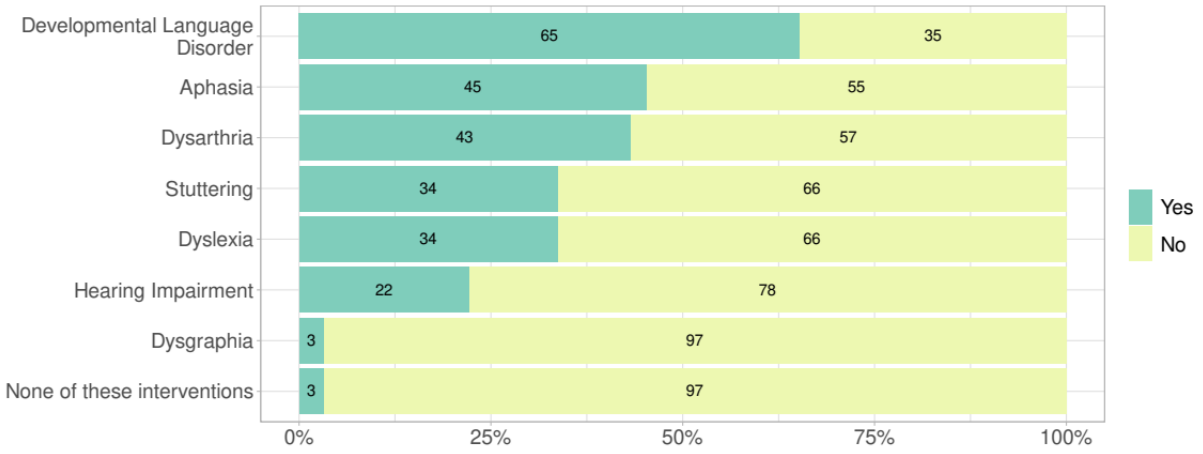

Figure 2. Use of methods based on music and rhythm in disorders for which theoretical frameworks report a beneficial role of such techniques. The numbers corresponding to "Yes" indicate the proportion of participants who selected each response option, displayed on the vertical axis.

One of the main goals when music and rhythm were used was to rehabilitate a specific symptom (69%), followed by the enhancement of communication in general (58%) (Figure 3). SLPs primarily chose to use music- and rhythm-based methods with their patients according to the pathology (81%), followed by their patients' ability to follow such interventions (69%) (Figure 4). Regarding the contexts in which music- and rhythm-based methods were used, 99% of SLPs used them in individual sessions and 25% in group sessions. Regarding the frequency of sessions dedicated to these methods, 67% answered 'occasionally', 9% 'every session', and 6% 'about one session in two'. The response 'other' was chosen by 17% of SLPs. During a session, music and rhythm were used for less than half of the session by 57% of respondents, for about half the session by 20%, for more than half of the session by 4%, and for the whole session by 6%. Thirteen percent of SLPs answered 'other'.

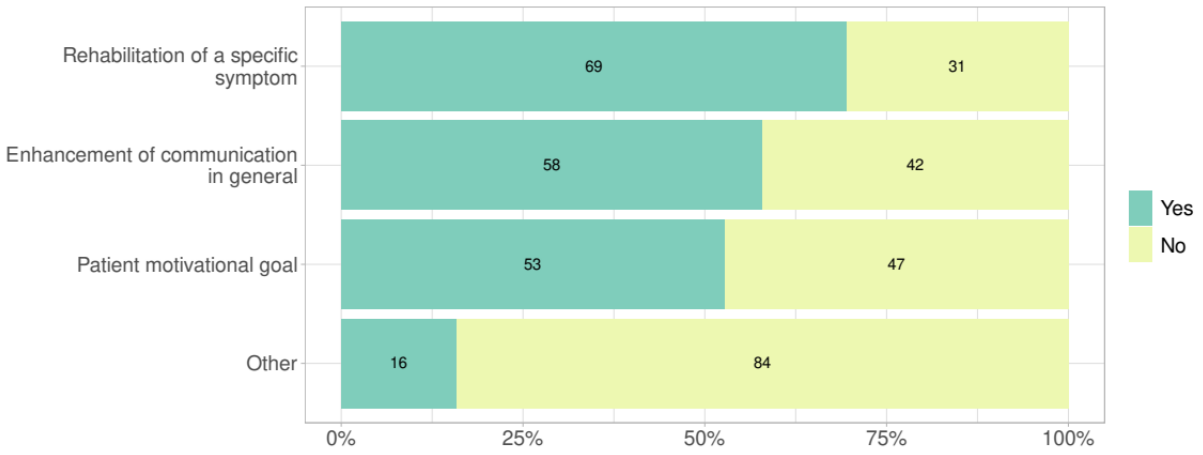

Figure 3. Therapeutic goals when using music- and rhythm-based interventions. The numbers corresponding to "Yes" indicate the proportion of participants who selected each response option, displayed on the vertical axis.

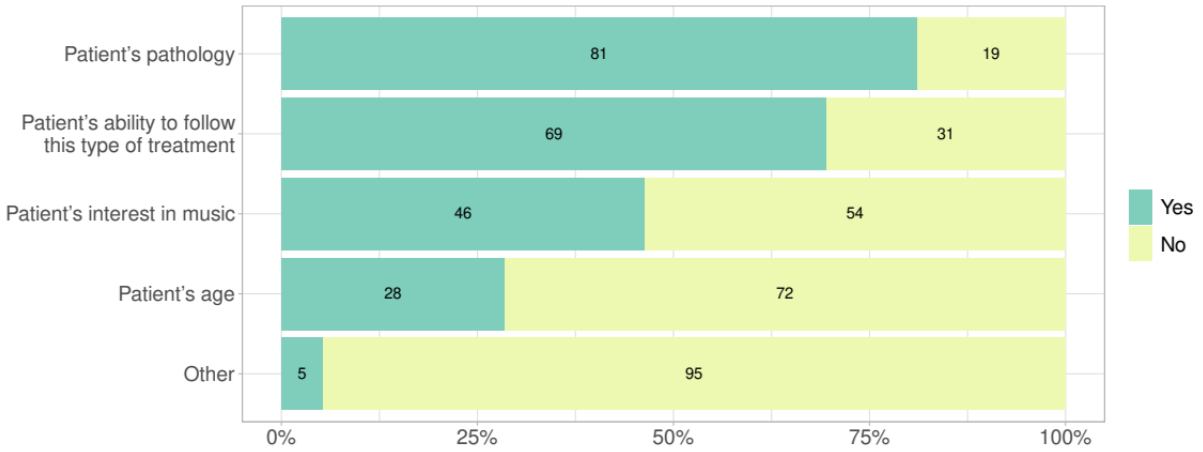

Figure 4. Criteria for using music- and rhythm-based interventions. The numbers corresponding to "Yes" indicate the proportion of participants who selected each response option, displayed on the vertical axis.

Regarding the perceived effect of such methods, 78% considered the effect to be positive, 20% could not estimate it, and 1% considered it neither positive nor negative. No participant felt that this type of intervention had a negative effect. One person did not answer. Participants who considered to effect to be positive (n=74) were asked to indicate the area of improvement. The most frequently reported area was the improvement of the targeted symptom (74%), followed by increased patient motivation to attend sessions (66%) (Figure 5).

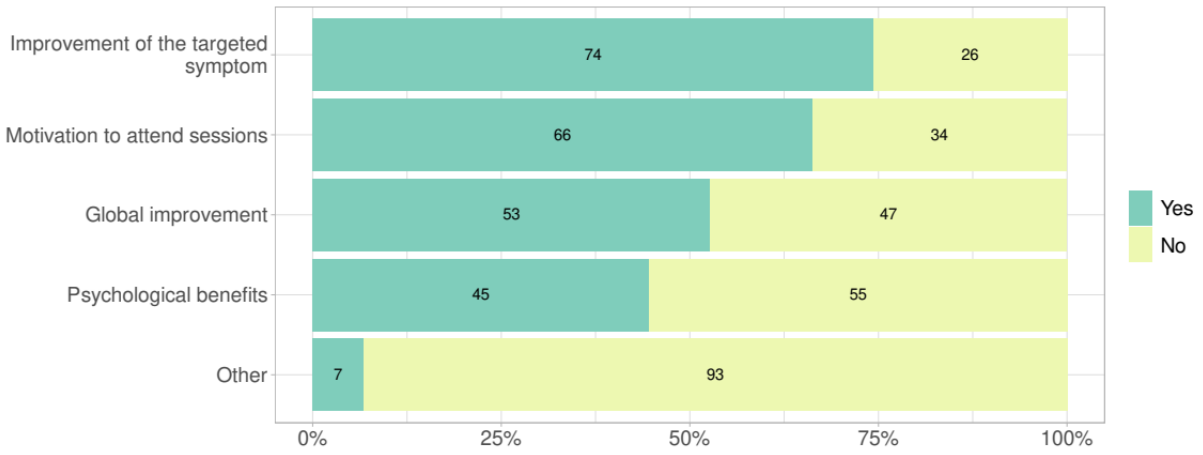

Figure 5. Perceived areas of improvement following use of music- and rhythm-based interventions. The numbers corresponding to "Yes" indicate the proportion of participants who selected each response option, displayed on the vertical axis**.**

3.3. Potential facilitators and barriers

This section addresses the second objective of the study, i.e., to identify factors associated with the use of music and rhythm in the clinical practices of SLPs. First, we asked SLPs what factors might facilitate the use of music- and rhythm-based methods in clinical practice. Only those using music and rhythm could answer this question (n=95). The results showed that the main facilitating factor was clinical knowledge and skills regarding music- and rhythm-based methods (68%), followed by musical skills (32%) and patients' requests (26%) (Figure 6A). One person did not answer this question. All participants were invited to answer a question about barriers. The most frequently reported barrier was the lack of training courses on these methods (61%), followed by the lack of information on their effectiveness (38%), and the lack of training or knowledge about them (33%) (Figure 6B). Five people did not answer this question.

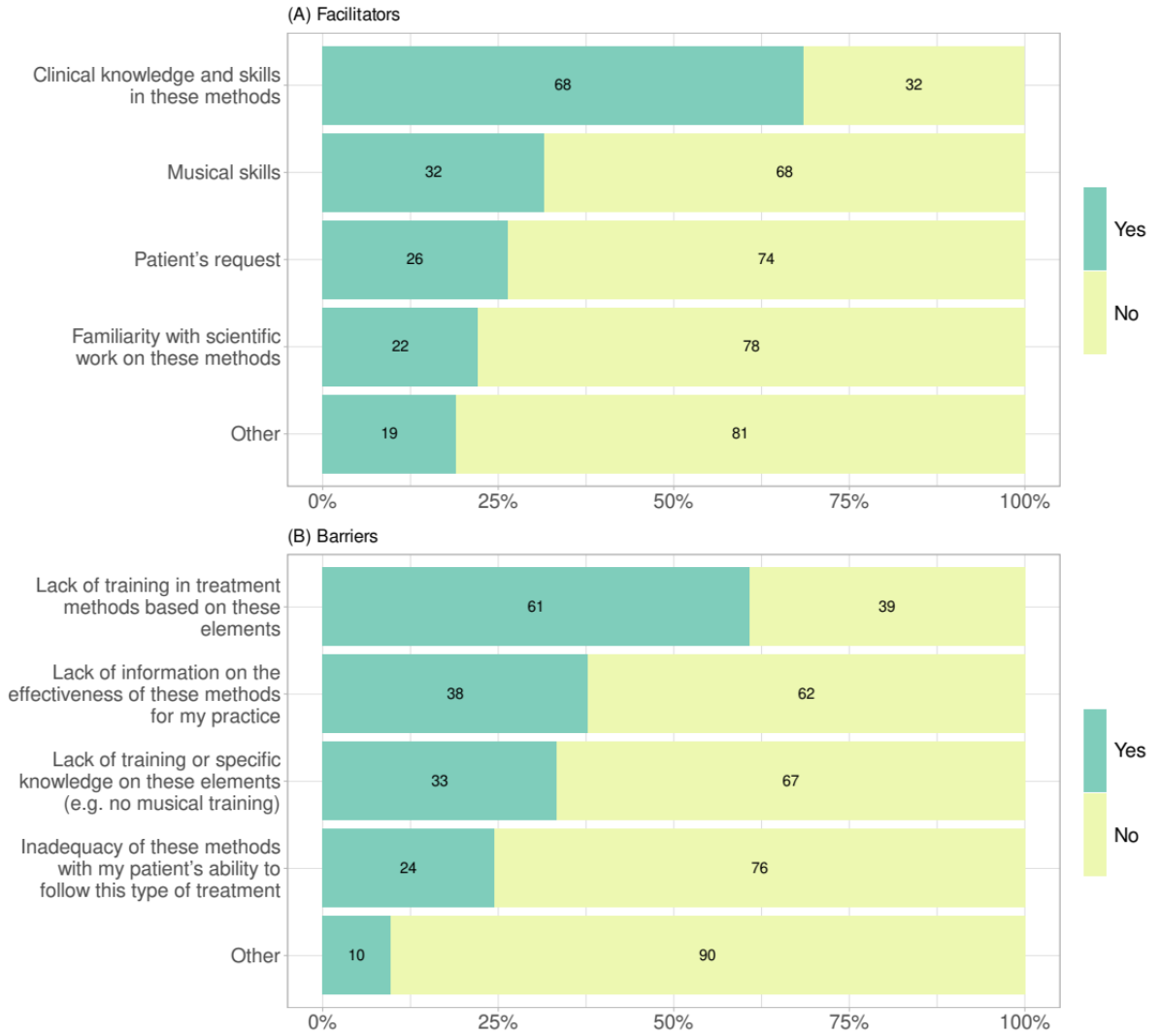

Figure 6. Facilitators (A) and barriers (B) for using music and rhythm in SLP practice. The numbers corresponding to "Yes" indicate the proportion of participants who selected each response option, displayed on the vertical axis.

We hypothesized that the use of music and rhythm in SLP interventions would be associated with the following: (1) formal extracurricular training in music and rhythm, (2) clinicians' personal music- and rhythm-related practices, (3) clinical training in music- and rhythm-based SLP methods, and (4) familiarity with the scientific literature on music- and rhythm-based rehabilitation approaches. Since questions addressing these aspects were mandatory, all SLPs provided responses (N=135). The results revealed significant associations between SLPs' clinical use of methods based on music and rhythm and the following factors: current music- and rhythm-related practices ($\chi^2(1) = 7.11$, $\phi = .23$, $p < .01$), clinical training in such methods ($\chi^2(1) = 10.95$, $\phi = .29$, $p < .001$), and familiarity with the scientific literature on rehabilitation

methods incorporating these features ($\tau = .22$, $z = 2.73$, $p < .01$). Participants with current music- and rhythm-related practices were more likely to incorporate music and rhythm into their clinical work than those without such practices (80% vs. 58%). Similarly, SLPs who had received clinical training in these methods reported using them more frequently than those without training (83% vs. 55%). Finally, SLPs who were more familiar with the scientific literature on music- and rhythm-based rehabilitation methods were more likely to use them in their everyday practice. However, no significant association was found between formal extracurricular musical training and the clinical use of these methods (73% vs. 56%, $\chi^2(1) = 1.81$, $\phi = .11$, $p = .18$).

4. Discussion

Although scientific research on the neurocognitive interplay between music and language has grown significantly, especially with regard to the mechanisms underlying the therapeutic benefits of music- and rhythm-based interventions, the everyday practices of SLPs remain underrepresented in this body of work. This study aimed to explore the use of these interventions in clinical practice and identify the barriers and facilitators affecting their implementation in SLP settings. The findings are discussed below.

4.1. Music and rhythm in SLPs' clinical practice

The key findings of this study concern whether and how SLPs use music and rhythm in their daily practice. Most SLPs reported using music- and rhythm-based techniques. The results show that music and rhythm are used to treat several pathologies and disorders mentioned in the main theoretical frameworks in cognitive psychology and cognitive neuroscience. These frameworks include OPERA (Patel, 2011, 2014), PRISM (Fiveash et al, 2021) and SEP (Fujii & Wan, 2014). SLPs reported using music- and rhythm-based techniques more frequently for developmental language disorder than for other disorders. As expected, they were also used to treat aphasia and dysarthria, though less consistently than in developmental language disorder. This disparity may stem from the fact that the respondents primarily treat children and adolescents rather than adults, as shown in Figure 1 and the Results section, which details evaluation and treatment practices over the past 12 months.

Despite the respondents' extensive experience with written language disorders, music- and rhythm-based interventions were used to treat dyslexia only half as much as for developmental language disorder. This disparity between spoken and written language disorders is unexpected and contradicts previous research. Studies have demonstrated a connection between developmental dyslexia and impaired temporal processing of auditory stimuli (e.g., Goswami, 2011; Ladányi et al, 2020). Furthermore, music- and rhythm-based methods have been considered effective for addressing reading difficulties. For instance, in their randomized controlled trial in Italian, Flaugnacco et al. (2015) compared the impact of music training and painting training (control condition) on the reading performance of children with dyslexia. The results showed that only music training enhanced reading accuracy for both text and pseudo-word reading. They also observed improved phonological abilities after music training. Similarly, Habib et al. (2016) demonstrated that their French-language *Cognitivo-Musical Training* program improved reading skills and phonological awareness in children with dyslexia. The systematic review by Cancer and Antonietti (2022) found that interventions based on music training yielded overall positive outcomes: while the impact on reading was positive but inconsistent, the impact on phonological abilities was consistently positive. The authors emphasized that integrating rhythmic components with decoding exercises may be the most effective approach for treating reading difficulties. However, our findings suggest that these evidence-based advances have not been widely integrated into routine clinical practice. This discrepancy may be due to a lack of training opportunities: only 55% of SLPs reported receiving instruction in music- and rhythm-based methods during their initial training or through post-certification continuing education. In addition, unlike well-established methods such as MIT, which is widely used in contexts like aphasia, no analogous music- or rhythm-based intervention exists for dyslexia. Importantly, few respondents reported being familiar with contemporary rehabilitation research on music- and rhythm-based interventions, which further underscores potential gaps in knowledge transfer.

Although most respondents reported having limited experience with stuttering, one third of SLPs still reported using music- and rhythm-based methods with patients who stutter. This is not surprising as simple rhythmic techniques such as using a metronome are well-established stuttering interventions (e.g., Trajkovski et al., 2009). Only a few SLPs reported incorporating music or rhythm into therapy for hearing impairment. This may reflect the composition of the

study sample, as very few participants had worked with hearing-impaired individuals in the past year. Future research should target SLPs who specialize in working with this population, as music-based approaches have been theorized and empirically demonstrated to promote speech and language development in children with hearing impairments or cochlear implants (Patel, 2011, 2014; Torppa & Huotilainen, 2019). Finally, music- and rhythm-based strategies were the least commonly used to treat dysgraphia in our study. Like hearing impairment, this area needs further investigation with SLPs who work with dysgraphia, as none of the participants reported recent experience with this population. This gap may also reflect the limited research on the relationship between music and writing in studies on writing mechanisms and dysgraphia (e.g., McCloskey & Rapp, 2017).

Regarding the context of use of music- and rhythm-based methods, most SLPs reported using these methods in individual sessions rather than group sessions. While the latter can be beneficial for patients (e.g., Pohl et al., 2020), they may be difficult to set up for logistical reasons, especially in private practice. Music- and rhythm-based methods are typically used for part of a session, not the entire session. Similarly, most SLPs use these methods only occasionally. They seem more likely to be used as a complement to other interventions.

SLPs reported that the primary goal of music- and rhythm-based interventions was symptom-specific improvement, and that the choice of method was largely dependent on the patient's clinical diagnosis. Most clinicians reported positive outcomes, citing symptom improvement as the most frequently observed benefit. Importantly, the results also show that music- and rhythm-based interventions may target and improve patient motivation. While our study did not investigate whether SLPs use these methods in a game-based format, it is noteworthy that serious games for music-based interventions are becoming more common (e.g., Zanto et al., 2024). Beyond addressing specific symptoms, they create engaging and motivating rehabilitation environments (Dalla Bella, 2022). Integrating rhythm and music components into game-based formats may further amplify patient motivation in SLP practice. This possibility warrants empirical investigation in future studies.

Finally, the results suggest that the current training landscape may pose a challenge. Little over half of the clinicians had received formal clinical training in music- and rhythm-based methods. The majority reported training focused specifically on rhythmic movement techniques. Interestingly, when SLPs were asked to describe their methods, one of the most

frequently reported techniques was the active incorporation of bodily movement into the intervention. Overall, there was a greater emphasis on rhythmic components than on broader musical applications. These findings underscore the need for specialized expertise and formal clinical training in music-based rehabilitation, as the absence of such training currently hinders the broader adoption of these interventions in clinical practice (Grau-Sánchez et al., 2022).

4.2. Potential facilitators and barriers in using music- and rhythm-based interventions

One of the aims of our study was to investigate the potential barriers and facilitators that influence the use of music- and rhythm-based methods in clinical practice. SLPs identified clinical knowledge and skills in these methods as the most critical facilitator, while the absence of such training was reported as the most significant barrier. A lack of evidence supporting the effectiveness of these methods also emerged as a barrier, though to a lesser extent than a lack of training. Beyond these perceived facilitators and barriers, we explored the association between applying these methods in daily practice and two distinct types of factors: professional factors such as clinical training in these methods and familiarity with the scientific literature about them, and personal factors such as formal extracurricular training or current engagement in music or rhythm activities, which extend beyond standard SLP training.

The results showed that SLPs who lacked clinical training in music and rhythm implementation were less likely to use these methods. Additionally, familiarity with the scientific literature on these methods was positively associated with their use. This suggests that limited exposure to research may reduce their adoption. These findings are consistent with SLPs' self-reported barriers and facilitators. Although prior extracurricular musical training showed no link to clinical use, SLPs currently engaged in such activities were more likely to integrate these methods. This aligns with the study by Ong et al. (2024), who emphasized collaborating with music professionals when SLPs lack personal expertise. However, unlike in the latter controlled experimental context, our study focused on real-world clinical practice, underscoring the idea that hands-on experience may lead more effectively to implementation.

In summary, our findings show that insufficient clinical training remains a key barrier to adopting music- and rhythm-based interventions in SLP practices. They also suggest that improving SLPs' familiarity with the scientific evidence supporting these methods and fostering their engagement in rhythm and music activities could be critical for implementing

such methods. If they are validated by further research, these insights may necessitate the rethinking of clinical training curricula to hone the skills of SLPs in using music- and rhythm-based techniques.

4.3. Clinical implications

To our knowledge, this is the first study to provide a comprehensive overview of how music and rhythm are integrated into the real-world clinical practices of SLPs, while also identifying the key facilitators and barriers to implementing these approaches. Our findings suggest that enhancing clinical training and knowledge of evidence-based applications is key to improving the use of music and rhythm in practice. This could be achieved through targeted training programs combining theoretical foundations with hands-on workshops. Improved training would help bridge the gap between theoretical knowledge and clinical practice based on music and rhythm. Importantly, our findings suggest that SLPs are more likely to use available and well-known protocols. Cueing strategies and MIT-based methods, which are well described in textbooks (e.g., Thaut & Hoemberg, 2014), were reported as being used by the respondents. Developing and teaching standardized protocols would thus be beneficial, particularity in domains where a gap between theoretical knowledge and clinical practice is identified, such as written language disorders. An initial step could be a collective effort by clinicians and researchers to create and disseminate clear, evidence-based guidelines for the use of music- and rhythm-based techniques in SLP practice. These guidelines would enable SLPs to apply current knowledge and best practices, while also allowing them to adapt their interventions to specific disorders, as well as to each patient's skills and motivation.

Furthermore, the results show that SLPs' personal engagement with musical activities affects how they apply these methods professionally. This suggests that building clinicians' confidence in their musical or rhythmic abilities could be beneficial. In this context, co-intervention models involving music professionals such as music therapists could broaden therapeutic opportunities and provide clinicians with more reassurance when using these techniques. While such collaborations exist in interventions for both adults and children (Ong et al., 2024; Burns et al., 2025), implementing them in private practice remains challenging. An alternative approach is redesigning clinical training curricula to better incorporate skill development in music and rhythm. SLPs may also benefit from clinical observation frameworks

or simplified decision-making tools to identify scenarios in which music and rhythm could be particularly effective. These tools could consider factors such as the types of disorders and neurocognitive mechanisms targeted by these interventions and motivational factors that enhance patient engagement in therapy. However, advancing the evidence-based implementation of these methods in real-world settings will require further research on their practical application.

Our findings suggest that SLPs frequently incorporate bodily movement into their music- and rhythm-based interventions. This observation aligns with theoretical frameworks that propose that a key mechanism behind the benefits of music and rhythm is their capacity to strengthen sensorimotor coupling (e.g., Patel, 2011, 2014; Fiveash et al., 2021). Since the aim of the present study was to investigate how non-verbal factors related to rhythm and music are used in SLP practice, the participants were asked not to consider singing when completing the questionnaire. Nevertheless, when asked to openly describe their methods related to rhythm and music, many mentioned incorporating singing-based methods. This may reflect a greater level of SLP training or awareness of such techniques or the perception that these approaches are more efficient, appropriate, or accessible than instrumental ones. Further research is needed to better understand the use of singing-based approaches compared to purely instrumental ones in SLP practice and the motivations underlying the choice of each approach. Because singing inherently relies on the musical element of rhythm, further investigation into singing combined with rhythm-based techniques (other than MIT) may also provide valuable insights.

4.4. Musical interventions and the issue of transfer

Our study was not designed to measure the effectiveness of music- and rhythm-based interventions, so we cannot offer recommendations in this area. However, this remains a critical concern. Two key issues warrant attention in future research. First, studies should investigate methodologies that SLPs could use to evaluate the impact of integrating music and rhythm into their daily practice. Second, although music- and rhythm-based approaches appear theoretically grounded and have shown promise in treating various pathologies, there are methodological limitations that must be addressed (Grau-Sánchez et al., 2022). Moreover, the strength of the evidence supporting their clinical application requires further research (Ree

et al., 2025). This issue extends beyond the field of SLP and concerns the relevance of music training as a tool for achieving therapeutic and educational benefits in clinical and academic settings. An important consideration in SLP practice is assessing the relative effectiveness of different approaches, such as singing-based vs. instrumental-only methods. In their study with neurologically impaired individuals, Cohen and Masse (1993) provided preliminary evidence that singing-based interventions had different therapeutic outcomes in terms of verbal intelligibility compared to rhythm-based interventions. It is crucial to determine whether the rhythmic elements of music are more effectively conveyed through singing rather than other means (see Stahl et al., 2011, for a discussion on this topic).

The use of music- and rhythm-based methods is motivated, at least in part, by the assumption that musical training improves cognitive abilities, implying a causal relationship between musical training and non-musical skills. It is noteworthy that there is an ongoing debate regarding the transfer assumption and its underlying nature (e.g., Schellenberg & Lima, 2024; Jamey et al.; 2024). To better address this issue within the context of the music–language relationship, although theoretical frameworks offer mechanistic explanations for transfer effects (Patel, 2011, 2014; Fujii & Wan, 2014; Fiveash et al., 2021), experimental studies must more rigorously disentangle the components of musical training to clarify the underlying neurocognitive mechanisms. In Alzheimer's disease patients, music-based interventions have been reported to stimulate social bonding (Matziorinis & Koelsch, 2022). These results highlight the role of emotional and social factors in understanding the benefits of music in clinical interventions. If the social effects of engaging with music are key, then interactive group sessions may be more effective. Conversely, if components such as temporal processing are central, then exercises targeting these components should take priority. Identifying the mechanisms at play is critical for optimizing clinical practice in speech-language pathology.

4.5. Limitations

This study has several limitations, some of which are related to the sample. First, the questionnaire was completed voluntarily, so respondents may have self-selected based on their interest in music- and rhythm-based methods, which could have biased the results. Second, around 80% of participants had completed at least one extracurricular training program related to music and rhythm, and slightly more than half of the SLPs were currently

playing an instrument, singing, or dancing. The SLPs who chose to participate may thus differ from those who did not complete the survey in their knowledge, consideration, and use of music and rhythm in interventions. Third, participants were asked not to consider singing when completing the questionnaire. However, in an open-ended question, some mentioned vocal techniques. As a result, some of our findings may have been influenced by this consideration. Fourth, the sample over-represents clinicians practicing in France, which limits the generalizability of the findings to other contexts. Fifth, the small sample size relative to the target population further constrains the robustness of the conclusions. Sixth, the questionnaire design may also have introduced constraints, as it did not include questions about how the perceived benefits of these methods were evaluated or about potential collaborations with other professionals (e.g., music therapists). Despite these limitations, this study provides a comprehensive exploration of how music and rhythm are integrated into real-world clinical practice.

5. Conclusion

This study explored the clinical practices of SLPs regarding the use of music- and rhythm-based interventions. The results show that SLPs use rhythm and music in interventions for various disorders, primarily to improve specific symptoms, and that the choice of method largely depends on the patient's clinical diagnosis. They also show how important it is to translate theoretical insights from scientific research into practical clinical tools. The findings also identified insufficient clinical training in these methods as a significant barrier to their adoption. In addition to clinical training, enhancing evidence-based knowledge by delving into the scientific literature and fostering personal engagement with music and rhythm may be essential to improve the integration of these approaches. However, the study did not examine the effectiveness of music- and rhythm-based interventions, the mechanisms underlying their potential benefits, or objective measurements of their implementation and evaluation by SLPs. Further research is thus needed to address these gaps and to develop reliable strategies that enable SLPs to evaluate the efficacy of such interventions within their routine clinical workflows.

Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Acknowledgments

We are grateful to the participants and to those who helped distribute the survey. We also thank those who reviewed the initial draft of the questionnaire. This work was supported by the French National Agency for Research (project ANR-21-CE28-0011).

References

Behaghel, E., & Zumbansen, A. (2022). Singing for the Rehabilitation of Acquired Neurogenic Communication Disorders: Continuing the Evidence Dialogue with a Survey of Current Practices in Speech-Language Pathology. *Healthcare*, *10*(6), Article 6. https://doi.org/10.3390/healthcare10061010

Burns, J., Keaveney, C., Nieto, N., O'Connor, R., & Moss, H. (2025). Collaborative Music Therapy and Speech-Language Pathology for Pediatric Acquired Communication Impairments: A Phenomenological International Perspective. *Music Therapy Perspectives*, *43*(1), miae026. https://doi.org/10.1093/mtp/miae026

Cancer, A., & Antonietti, A. (2022). Music-based and auditory-based interventions for reading difficulties: A literature review. *Heliyon*, *8*(4). https://doi.org/10.1016/j.heliyon.2022.e09293

Cason, N., Hidalgo, C., Isoard, F., Roman, S., & Schön, D. (2015). Rhythmic priming enhances speech production abilities: Evidence from prelingually deaf children. *Neuropsychology*, *29*(1), 102–107. https://doi.org/10.1037/neu0000115

Cohen, N. S., & Masse, R. (1993). The application of singing and rhythmic instruction as a therapeutic intervention for persons with neurogenic communication disorders. *Journal of Music Therapy*, *30*(2), 81-99.

Colling, L. J., Noble, H. L., & Goswami, U. (2017). Neural Entrainment and Sensorimotor Synchronization to the Beat in Children with Developmental Dyslexia: An EEG Study. *Frontiers in Neuroscience*, *11*, 360. https://doi.org/10.3389/fnins.2017.00360

Daigmorte, C., Tallet, J., & Astésano, C. (2022). On the foundations of rhythm-based methods in Speech Therapy. *Speech Prosody 2022*, 47-51. https://doi.org/10.21437/SpeechProsody.2022-10

Dalla Bella, S. (2022). Rhythmic serious games as an inclusive tool for music-based interventions. *Annals of the New York Academy of Sciences*, *1517*(1), 15–24. https://doi.org/10.1111/nyas.14878

Fiveash, A., Bedoin, N., Gordon, R. L., & Tillmann, B. (2021). Processing rhythm in speech and music : Shared mechanisms and implications for developmental speech and language disorders. *Neuropsychology*, *35*(8), 771-791. https://doi.org/10.1037/neu0000766

Flaugnacco, E., Lopez, L., Terribili, C., Montico, M., Zoia, S., & Schön, D. (2015). Music training increases phonological awareness and reading skills in developmental dyslexia: A randomized control trial. *PloS one*, *10*(9), e0138715. https://doi.org/10.1371/journal.pone.0138715

Frankford, S. A., Murray, E. S. H., Masapollo, M., Cai, S., Tourville, J. A., Nieto-Castañón, A., & Guenther, F. H. (2021). The Neural Circuitry Underlying the « Rhythm Effect » in Stuttering. *Journal of Speech, Language & Hearing Research*, *64*, 2325-2346. https://doi.org/10.1044/2021_JSLHR-20-00328

Fujii, S., & Wan, C. Y. (2014). The Role of Rhythm in Speech and Language Rehabilitation : The SEP Hypothesis. *Frontiers in Human Neuroscience*, *8*, 777. https://doi.org/10.3389/fnhum.2014.00777

Gehlbach, H., & Artino, A. R. (2018). The Survey Checklist (Manifesto). *Academic Medicine: Journal of the Association of American Medical Colleges*, *93*(3), 360-366. https://doi.org/10.1097/ACM.0000000000002083

Goswami, U. (2011). A temporal sampling framework for developmental dyslexia. *Trends in Cognitive Sciences*, *15*(1), 3–10. https://doi.org/10.1016/j.tics.2010.10.001

Grau-Sánchez, J., Jamey, K., Paraskevopoulos, E., Dalla Bella, S., Gold, C., Schlaug, G., Belleville, S., Rodríguez-Fornells, A., Hackney, M. E., & Särkämö, T. (2022). Putting music to trial: Consensus on key methodological challenges investigating music-based rehabilitation. *Annals of the New York Academy of Sciences*, *1518*(1), 12–24. https://doi.org/10.1111/nyas.14892

Habib, M., Lardy, C., Desiles, T., Commeiras, C., Chobert, J., & Besson, M. (2016). Music and Dyslexia: A New Musical Training Method to Improve Reading and Related Disorders. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.00026

Jamey, K., Foster, N. E. V., Hyde, K. L., & Dalla Bella, S. (2024). Does music training improve inhibition control in children? A systematic review and meta-analysis. Cognition, 252, 105913. https://doi.org/10.1016/j.cognition.2024.105913

Johnson, M. A., Coles, H., Keough, L., King, B., & Reed, M. (2019). Co-Delivered Integrative Music and Language Therapy : Positive Outcomes Through Music Therapy and Speech-

Language Pathology Collaboration. *Perspectives of the ASHA Special Interest Groups*, *4*(2), 261-268. https://doi.org/10.1044/2019_PERS-SIG2-2018-0006

Kotz, S. A., & Gunter, T. C. (2015). Can rhythmic auditory cuing remediate language-related deficits in Parkinson's disease? *Annals of the New York Academy of Sciences*, *1337*(1), 62-68. https://doi.org/10.1111/nyas.12657

Ladányi, E., Persici, V., Fiveash, A., Tillmann, B., & Gordon, R. L. (2020). Is atypical rhythm a risk factor for developmental speech and language disorders?. *Wiley Interdisciplinary Reviews: Cognitive Science*, *11*(5), e1528. https://doi.org/10.1002/wcs.1528

Ladányi, E., Lukács, Á., & Gervain, J. (2021). Does rhythmic priming improve grammatical processing in Hungarian-speaking children with and without developmental language disorder? *Developmental Science*, *24*(6), e13112. https://doi.org/10.1111/desc.13112

McCloskey, M., & Rapp, B. (2017). Developmental dysgraphia: An overview and framework for research. *Cognitive Neuropsychology*, *34*(3–4), 65–82. https://doi.org/10.1080/02643294.2017.1369016

Mainka, S., & Mallien, G. (2014). Rhythmic speech cueing (RSC). In M. H. Thaut & V. Hoemberg (Eds.), *Handbook of neurologic music therapy* (pp. 150–160). Oxford University Press.

Matziorinis, A. M., & Koelsch, S. (2022). The promise of music therapy for Alzheimer's disease: A review. *Annals of the New York Academy of Sciences*, *1516*(1), 11-17. https://doi.org/10.1111/nyas.14864

Ong, A., Namasivayam-MacDonald, A., Kim, S., & Werden Abrams, S. (2024). The use of music and music-related elements in speech-language therapy interventions for adults with neurogenic communication impairments: A scoping review. *International Journal of Language & Communication Disorders*, *59*(6), 2632-2654. https://doi.org/10.1111/1460-6984.13104

Patel, A. D. (2011). Why would Musical Training Benefit the Neural Encoding of Speech? The OPERA Hypothesis. *Frontiers in Psychology*, *2*, 142. https://doi.org/10.3389/fpsyg.2011.00142

Patel, A. D. (2014). Can nonlinguistic musical training change the way the brain processes speech? The expanded OPERA hypothesis. *Hearing Research*, *308*, 98-108. https://doi.org/10.1016/j.heares.2013.08.011

Pohl, P., Wressle, E., Lundin, F., Enthoven, P., & Dizdar, N. (2020). Group-based music intervention in Parkinson's disease – findings from a mixed-methods study. *Clinical Rehabilitation*, *34*(4), 533–544. https://doi.org/10.1177/0269215520907669

Ree, A., Denmon, C., Tibus, E. O., Mayes, A., Weatherill, M., Auerbach, A., & Rodriguez, A. D. (2025). Research on music-based interventions for aphasia: a scoping review. *Aphasiology*, 1-28. https://doi.org/10.1080/02687038.2025.2451670

Schellenberg, E. G., & Lima, C. F. (2024). Music training and nonmusical abilities. *Annual Review of Psychology*, *75*(1), 87-128. https://doi.org/10.1146/annurev-psych-032323-051354

Sihvonen, A. J., Särkämö, T., Leo, V., Tervaniemi, M., Altenmüller, E., & Soinila, S. (2017). Music-based interventions in neurological rehabilitation. *The Lancet. Neurology*, *16*(8), 648-660. https://doi.org/10.1016/S1474-4422(17)30168-0

Sparks, R. W., & Holland, A. L. (1976). Method: Melodic Intonation Therapy for Aphasia. *Journal of Speech and Hearing Disorders*, *41*(3), 287-297. https://doi.org/10.1044/jshd.4103.287

Stahl, B., Kotz, S. A., Henseler, I., Turner, R., & Geyer, S. (2011). Rhythm in disguise: Why singing may not hold the key to recovery from aphasia. *Brain*, *134*(10), 3083–3093. https://doi.org/10.1093/brain/awr240

Thaut, M., & Hoemberg, V. (2014). *Handbook of Neurologic Music Therapy*. Oxford University Press.

Torppa, R., & Huotilainen, M. (2019). Why and how music can be used to rehabilitate and develop speech and language skills in hearing-impaired children. *Hearing research*, *380*, 108-122. https://doi.org/10.1016/j.heares.2019.06.003

Trajkovski, N., Andrews, C., Onslow, M., Packman, A., O'Brian, S., & Menzies, R. (2009). Using syllable-timed speech to treat preschool children who stutter: A multiple baseline experiment. *Journal of fluency disorders*, *34*(1), 1-10. https://doi.org/10.1016/j.jfludis.2009.01.001

Van Nuffelen, G., De Bodt, M., Vanderwegen, J., Van de Heyning, P., & Wuyts, F. (2010). Effect of rate control on speech production and intelligibility in dysarthria. *Folia Phoniatrica et*

Logopaedica: Official Organ of the International Association of Logopedics and Phoniatrics (IALP), *62*(3), 110-119. https://doi.org/10.1159/000287209

Van Tellingen, M., Hurkmans, J., Terband, H., Jonkers, R., & Maassen, B. (2023). Music and musical elements in the treatment of childhood speech sound disorders: A systematic review of the literature. *International Journal of Speech-Language Pathology*, *25*(4), 549-565. https://doi.org/10.1080/17549507.2022.2097310

Véron-Delor, L., Pinto, S., Eusebio, A., Azulay, J. P., Witjas, T., Velay, J. L., & Danna, J. (2020). Musical sonification improves motor control in Parkinson's disease: a proof of concept with handwriting. *Annals of the New York Academy of Sciences*, *1465*(1), 132-145. https://doi.org/10.1111/nyas.14252

Zanto, T. P., Giannakopoulou, A., Gallen, C. L., Ostrand, A. E., Younger, J. W., Anguera-Singla, R., ... & Gazzaley, A. (2024). Digital rhythm training improves reading fluency in children. *Developmental Science*, *27*(3), e13473. https://doi.org/10.1111/desc.1347

Troisième Partie

Discussion Générale

Chapitre 1
Synthèse transversale

L'objectif principal de cette thèse était d'examiner dans quelle mesure un rythme auditif externe peut influencer la production de la parole dans la MP, en particulier ses dimensions temporelles et prosodiques. Cette question s'inscrivait dans une double réalité. D'une part, la MP constitue une pathologie dans laquelle les troubles de la parole sont fréquents, fonctionnellement importants, et encore insuffisamment compris dans leur dimension prosodique. D'autre part, les interventions et paradigmes fondés sur le rythme offrent une piste intéressante en raison du rôle central des régularités temporelles dans la perception, l'anticipation et l'action.

La première partie de ce manuscrit posait ainsi plusieurs constats. Le premier était que la dysarthrie hypokinétique ne peut être réduite à une atteinte vocale ou articulatoire isolée. Elle implique aussi une altération de l'organisation temporelle et prosodique de la parole, avec des perturbations de la modulation mélodique, de l'accentuation, du débit et des pauses. Le deuxième constat était que, si les approches rythmiques ont été largement étudiées dans la MP pour la marche et d'autres comportements moteurs, leur transposition à la parole reste beaucoup moins documentée. Le troisième était que les cadres théoriques disponibles permettent de formuler des hypothèses mécanistiques, mais ne décrivent pas encore précisément comment un rythme externe pourrait modifier la production de la parole. C'est dans cette perspective que deux cadres ont été privilégiés : PRISM (Fiveash et al., 2021), pour penser les mécanismes partagés entre rythme, musique et parole, et le modèle neurocomputationnel DIVA/GODIVA, pour situer l'effet potentiel du rythme dans l'architecture de la production parlée (Bohland et al., 2010 ; Guenther, 2016 ; Miller & Guenther, 2021 ; Tourville & Guenther, 2011).

Les travaux présentés dans la deuxième partie de ce manuscrit répondent à ces enjeux de manière progressive. La méta-analyse consacrée à la prosodie dans la MP a d'abord permis d'identifier les dimensions acoustiques les plus robustement altérées dans cette population. L'étude expérimentale menée auprès des populations cliniques a ensuite examiné l'effet

d'amorces rythmiques sur l'initiation, le débit et des mesures rythmiques de la parole, puis sur des dimensions prosodiques plus ciblées, notamment la variabilité de F0 et les pauses. Le chapitre suivant a prolongé cette réflexion auprès d'adultes tout-venants, en examinant plus finement l'effet de la régularité, de la structure de l'amorce, de la structure prosodique de la phrase et des compétences de synchronisation sensorimotrice sur la production de parole. Enfin, le dernier chapitre, consacré aux pratiques orthophoniques, a replacé ces questions dans une perspective clinique, en examinant la manière dont la musique et le rythme sont actuellement connus, utilisés et évalués par les orthophonistes.

Identifier les dimensions prosodiques les plus robustement altérées dans la maladie de Parkinson

Le premier objectif de la thèse était d'identifier les dimensions temporelles et prosodiques les plus robustement altérées dans la MP. Les descriptions cliniques classiques évoquent une réduction de la variation de hauteur, une monotonie d'intensité, une réduction de l'accentuation, des pauses inappropriées, un débit variable ou encore des accélérations brèves. Toutefois, ces dimensions ne sont pas toujours confirmées avec la même robustesse par les mesures acoustiques.

La méta-analyse présentée dans le premier chapitre a permis de clarifier ce point. Les résultats montrent que les deux paramètres les plus solidement associés à la MP sont la réduction de la variabilité de F0 et l'augmentation de la durée des pauses. En revanche, les preuves sont plus limitées pour la variabilité d'intensité et le débit de parole, et moins convaincantes pour le débit articulatoire ou le ratio de pauses (Fumel et al., 2024).

La réduction de la variabilité de F0 s'inscrit dans les descriptions classiques de la parole parkinsonienne, souvent perçue comme plus monotone ou moins expressive. Elle renvoie à une difficulté à moduler la hauteur de manière suffisamment contrastée pour soutenir l'intonation, l'accentuation ou certaines fonctions discursives. L'augmentation de la durée des pauses relève d'une autre dimension : elle touche l'organisation temporelle de l'énoncé, la segmentation de la parole et possiblement la planification ou la transition entre unités. Ces deux dimensions partagent une composante temporelle, mais elles ne reflètent pas nécessairement les mêmes niveaux de contrôle. La première concerne davantage la

dynamique mélodique et le contrôle laryngé ; la seconde engage le phrasé, la continuité de la production, la respiration, la planification et le passage d'une unité à l'autre. Les modèles actuels de production de la parole ne permettent pas encore de rendre compte finement de ces dimensions sur le plan neurocognitif, bien que de tels modèles soient indispensables pour mieux comprendre la dysprosodie dans la MP.

Un effet de l'amorçage rythmique principalement porté par l'initiation de la parole

La thèse avait également pour objectif de déterminer si l'exposition à une amorce rythmique auditive modifie effectivement la production de parole ultérieure, et d'identifier les dimensions les plus sensibles à cette modulation. Les résultats concernent principalement le temps d'initiation. Dans l'étude menée auprès des participants contrôles, des personnes atteintes de la MP et des personnes atteintes d'ASC, ainsi que dans l'étude auprès d'adultes tout-venants, les amorces régulières ont raccourci le temps d'initiation de la parole. Ce résultat soutient l'hypothèse selon laquelle un rythme régulier peut faciliter la préparation temporelle d'une réponse vocale.

Ce résultat est cohérent avec la DAT, selon laquelle les régularités temporelles externes peuvent orienter l'attention vers des moments attendus dans le temps et faciliter le traitement ou la préparation d'événements subséquents (Jones & Boltz, 1989 ; Large & Jones, 1999 ; Jones, 2019). Il rejoint également le cadre PRISM, qui propose que les liens entre rythme, musique et parole reposent notamment sur le traitement auditif précis, la synchronisation aux régularités temporelles et le couplage sensorimoteur (Fiveash et al., 2021). Dans cette perspective, une amorce régulière peut créer un contexte temporel prédictible, susceptible de préparer le système à l'action qui suit.

Ce point prend un relief particulier dans la MP. La première partie de ce manuscrit rappelait que les ganglions de la base sont impliqués dans l'initiation, la sélection et la stabilisation de l'action. Dans le modèle DIVA/GODIVA, ces circuits participent au déclenchement et au maintien des programmes moteurs de parole, en interaction avec les régions prémotrices et motrices, notamment la SMA et la pré-SMA (Bohland et al., 2010 ; Guenther, 2016 ; Tourville & Guenther, 2011). L'application récente de ces modèles à la MP proposée par Manes et al. (2024) permet d'interpréter les troubles de la parole comme

résultant, au moins en partie, d'une difficulté à activer, maintenir et calibrer les programmes moteurs de parole. Dans cette perspective, le fait que les amorces régulières facilitent le démarrage de la parole suggère qu'un repère temporel externe peut soutenir le déclenchement de la séquence parlée.

L'effet observé chez les personnes atteintes de la MP est toutefois moins marqué que chez les contrôles. Ce point est important, car il nuance l'hypothèse d'une compensation simple. Les personnes atteintes de la MP ne sont pas incapables d'utiliser le rythme externe ; au contraire, elles en bénéficient. Mais ce bénéfice est atténué, ce qui suggère que les circuits nécessaires à l'exploitation de cette structure temporelle restent partiellement fragilisés. La régularité fournit donc un appui, mais cet appui ne restaure pas entièrement les conditions de préparation observées chez les participants contrôles.

Cette interprétation est renforcée par les résultats obtenus dans une deuxième expérience auprès d'adultes tout-venants (chapitre 4, deuxième partie). Les amorces régulières étaient également associées à des temps d'initiation plus courts que les amorces irrégulières. Ce résultat montre que l'effet de l'amorçage rythmique sur l'entrée en parole n'est pas propre aux populations cliniques ; il s'observe également chez des locuteurs sans pathologie neurologique. Il suggère donc que l'amorce régulière peut agir sur un mécanisme général de préparation temporelle de la parole, dont l'expression peut ensuite varier selon le profil neurofonctionnel des participants.

Les effets sur les autres dimensions de parole sont en revanche moins constants. Le débit de parole n'est pas modifié de manière homogène par les amorces régulières. Dans l'étude menée auprès des populations cliniques, les personnes atteintes d'ASC ralentissent leur parole après amorce régulière, tandis que les personnes atteintes de la MP et les contrôles ne montrent pas d'effet comparable. Dans le chapitre mené auprès de tout-venants, le débit est surtout déterminé par la structure prosodique de la phrase et non par l'amorce.

Cette conclusion répond de manière nuancée à l'hypothèse initiale d'un effet de l'amorçage sur la production. Oui, l'amorçage rythmique modifie la parole mais il la modifie d'abord au niveau de la préparation temporelle et du déclenchement. Il ne semble pas produire une reconfiguration générale de l'ensemble de la production une fois la parole lancée.

Une modulation prosodique sélective plutôt qu'un effet global

Un autre enjeu de la thèse était de déterminer si l'amorçage rythmique agit uniquement sur des paramètres globaux, comme l'initiation ou le débit, ou s'il peut également s'étendre à d'autres dimensions prosodiques notamment la variabilité de F0 et les pauses.

Chez les personnes atteintes de la MP, la variabilité de F0 est plus élevée après amorce régulière qu'après amorce irrégulière. Ce résultat est intéressant, car il concerne précisément l'un des marqueurs les plus robustes identifiés dans la méta-analyse (Fumel et al., 2024). La réduction de la variabilité de F0 constitue en effet une caractéristique importante de la dysprosodie parkinsonienne. Le fait qu'une amorce régulière soit associée à une variabilité mélodique plus élevée que l'amorce irrégulière suggère que cette dimension peut être sensible au contexte rythmique précédant la parole.

Il convient toutefois de ne pas surinterpréter ce résultat. L'effet oppose la condition régulière à la condition irrégulière ; il ne démontre pas une augmentation de la variabilité de F0 par rapport à la condition sans amorce. Il ne permet donc pas de conclure à une normalisation globale de la variabilité mélodique de la parole. Cette prudence est d'autant plus nécessaire que, dans le présent échantillon, les personnes atteintes de la MP ne se distinguaient pas clairement des participants contrôles sur cette mesure. L'effet observé doit donc être interprété comme une modulation conditionnelle de la dynamique mélodique par la structure de l'amorce, plutôt que comme la correction d'un déficit prosodique objectivé dans cette étude. Il indique néanmoins que la structure de l'amorce entendue avant la production peut moduler certains aspects de la dynamique mélodique chez les personnes atteintes de la MP. Cet effet reste circonscrit, mais il est important, car il suggère que l'influence de l'amorce ne se limite pas au déclenchement de la parole et peut concerner, plus ponctuellement, certaines dimensions prosodiques.

Les résultats sur les pauses conduisent à une autre lecture. Alors que la durée des pauses apparaît comme un marqueur robuste de la MP dans la méta-analyse, l'étude expérimentale ne met pas en évidence d'effet clair de l'amorçage sur les pauses chez les personnes atteintes de la MP. Les différences les plus robustes concernent plutôt le facteur de groupe : les personnes atteintes d'ASC produisent des pauses plus longues et réalisent plus fréquemment certaines pauses attendues. Ce profil ne traduit donc pas un effet spécifique de l'amorçage, mais plutôt une organisation temporelle particulière de la parole dans l'ASC,

marquée par une segmentation plus forte de l'énoncé et par un coût accru aux frontières prosodiques.

Ces résultats invitent donc à distinguer les effets de l'amorçage des différences liées aux profils cliniques. Alors que la variabilité de F0 chez les personnes atteintes de la MP semble pouvoir être ponctuellement modulée par la structure de l'amorce, les mesures de pauses apparaissent davantage liées aux caractéristiques propres des groupes, en particulier chez les personnes atteintes d'ASC. Les données ne soutiennent donc pas l'existence d'un effet prosodique global de l'amorçage.

Apport de la comparaison entre populations cliniques

Un autre objectif de la thèse était de comparer les personnes atteintes de la MP, les personnes atteintes d'ASC et les participants contrôles, afin de mieux contraindre l'interprétation mécanistique des effets observés. Cette comparaison s'inscrivait dans un cadre général distinguant les contributions respectives, mais non exclusives, des ganglions de la base et du cervelet au traitement temporel, à la préparation de l'action et au contrôle sensorimoteur (Grahn & Brett, 2009 ; Teki et al., 2011 ; Kotz & Schwartze, 2016). Elle répondait également à une question centrale pour les interventions basées sur le rythme dans la MP : celle de savoir dans quelle mesure un repère temporel externe peut soutenir ou compenser certains processus de timing, d'initiation ou de séquençage moteur altérés dans cette pathologie (Nombela et al., 2013 ; Dalla Bella et al., 2015 ; Ghai et al., 2018 ; Puyjarinet et al., 2019).

Les résultats montrent que les deux populations cliniques partagent une sensibilité de l'initiation aux amorces régulières. Chez les personnes atteintes de la MP comme chez les personnes atteintes d'ASC, les amorces régulières sont associées à une réduction du temps d'initiation. Toutefois, cet effet apparaît crédiblement moins marqué chez les personnes atteintes de la MP que chez les participants contrôles. Cela suggère que le rythme externe peut faciliter l'entrée dans la parole chez les personnes atteintes de la MP, mais de manière plus limitée que chez les locuteurs sans pathologie. Une modulation ponctuelle de la variabilité de F0 selon la structure de l'amorce est également observée chez les personnes atteintes de la MP, ce qui indique que certaines dimensions mélodiques de la parole

parkinsonienne peuvent rester sensibles aux propriétés temporelles du contexte auditif préalable.

Chez les personnes atteintes d'ASC, l'effet sur l'initiation est également présent, mais leur profil se distingue surtout par des effets touchant le déroulement temporel de l'énoncé. Les amorces régulières sont associées à un ralentissement du débit et à une modulation de certaines métriques syllabiques, ce qui suggère que l'amorçage peut influencer l'organisation temporelle de la parole au-delà du seul déclenchement de la réponse. En revanche, les résultats observés sur les pauses semblent relever davantage du profil pathologique que d'un effet direct de l'amorce. Les personnes atteintes d'ASC produisent des pauses plus longues et réalisent plus fréquemment certaines pauses attendues, ce qui évoque une segmentation plus marquée de l'énoncé et une temporalité plus coûteuse aux frontières prosodiques.

Ainsi, la comparaison entre populations suggère que l'amorçage rythmique régulier facilite l'initiation de la parole dans les deux groupes cliniques, mais que ses effets ne s'organisent pas de la même manière selon le profil pathologique. Dans la MP, le bénéfice apparaît surtout au moment de l'entrée dans la parole, avec un effet crédible mais moins marqué que chez les participants contrôles. Les autres effets observés restent plus ponctuels et concernent certaines dimensions acoustiques ou prosodiques, notamment la variabilité consonantique et la variabilité de F0, pour lesquelles l'amorce irrégulière semble constituer un contexte moins favorable que l'amorce régulière. Chez les personnes atteintes d'ASC, l'effet sur l'initiation est également présent, mais il s'accompagne d'effets sur le déroulement temporel de l'énoncé, en particulier sur le débit et certaines métriques syllabiques. Les résultats concernant les pauses doivent toutefois être distingués de ces effets d'amorçage : ils semblent davantage refléter les caractéristiques générales de la parole cérébelleuse, marquée par des contraintes de coordination et de segmentation.

Cette comparaison ne revient pas à opposer frontalement les ganglions de la base et le cervelet. Ces structures appartiennent à des réseaux interconnectés, et la parole repose sur leurs interactions avec des régions corticales auditives, prémotrices, motrices et associatives. Les résultats invitent plutôt à considérer que les contraintes propres à chaque pathologie peuvent orienter différemment l'utilisation d'un même contexte rythmique. Ainsi, le rythme

régulier peut moduler certaines composantes de la production, selon le niveau de parole considéré et selon le profil neurologique des locuteurs.

Synchronisation sensorimotrice et variabilité interindividuelle

Ce travail visait également à examiner dans quelle mesure les effets de l'amorçage rythmique peuvent être reliés à certaines différences individuelles, notamment les compétences de synchronisation sensorimotrice et le profil musical des participants. Cet objectif est important, car les résultats des chapitres précédents montrent que l'effet du rythme n'est ni homogène ni systématique. Il dépend des propriétés des amorces, des caractéristiques de la phrase, de la population étudiée, et probablement aussi des ressources individuelles permettant d'extraire et d'exploiter une structure temporelle.

Le chapitre mené auprès de tout-venants apporte ici un éclairage complémentaire. En plus de l'effet de régularité observé sur le temps d'initiation, ce chapitre examinait les relations entre les variables de parole et plusieurs mesures de compétences rythmiques et musicales : consistance et précision aux tâches de tapping, seuils issus du BST de la H-BAT (Fujii & Schlaug, 2013) et score global de musicalité. Ces analyses avaient un statut exploratoire, mais elles permettaient de tester une hypothèse importante issue du cadre PRISM : les effets d'un rythme perçu sur la parole devraient dépendre, au moins en partie, de la capacité du participant à extraire une régularité temporelle et à l'utiliser pour faciliter l'action (Fiveash et al., 2021).

Les résultats des corrélations restent limités. Une seule association crédible a été observée : la consistance dans la tâche de tapping avec accentuation tous les deux temps était associée au bénéfice concernant le temps d'initiation, avec une corrélation négative faible à modérée. Autrement dit, les participants les plus consistants dans cette tâche tendaient à montrer un bénéfice plus marqué de l'amorce régulière par rapport à l'amorce irrégulière sur le temps d'initiation, dans une des conditions de l'étude. Ce résultat est compatible avec l'idée que de meilleures capacités de synchronisation peuvent favoriser l'exploitation d'une structure rythmique pour préparer la parole.

Cependant, cette association demeure isolée. Aucune relation robuste n'a été observée entre le score global de musicalité et les variables de parole, ni entre les autres indices rythmiques et les effets comportementaux. Il serait donc excessif de conclure que les

compétences rythmiques prédisent globalement la sensibilité à l'amorçage. Les résultats suggèrent plutôt une relation possible, mais faible et dépendante de la mesure considérée. La consistance dans une tâche de tapping accentué pourrait refléter une compétence particulièrement pertinente pour la tâche, parce qu'elle implique non seulement une synchronisation régulière, mais aussi la prise en compte d'une structure accentuelle. Cette dimension se rapproche davantage de la structure des amorces utilisées dans l'expérience qu'un simple maintien de tempo ou qu'un score global de musicalité.

Ce résultat ne permet pas d'identifier un profil clair de « répondeur » à l'amorçage rythmique, mais il confirme l'intérêt de prendre en compte la variabilité interindividuelle. Les compétences rythmiques et sensorimotrices ne doivent pas être considérées comme des variables périphériques : elles constituent probablement l'un des niveaux nécessaires pour comprendre pourquoi une même amorce peut faciliter la parole chez certains participants, rester sans effet chez d'autres, ou agir sur des dimensions différentes de la production. Cette interprétation rejoint les travaux montrant que les capacités de perception et de production du beat varient fortement entre individus et peuvent être dissociées selon les dimensions évaluées (Fujii & Schlaug, 2013 ; Dalla Bella et al., 2017). Elle s'inscrit également dans une littérature plus large reliant les capacités de synchronisation au beat à d'autres dimensions perceptives, cognitives ou linguistiques (Tierney & Kraus, 2013), ainsi qu'aux capacités de prédiction temporelle et de synchronisation sensorimotrice (Nozaradan et al., 2016).

Articuler PRISM et DIVA/GODIVA

L'un des intérêts théoriques de cette thèse est de mettre en regard deux cadres qui répondent à des questions différentes. PRISM permet de comprendre comment une structure rythmique perçue peut être encodée, suivie et transformée en ressource temporelle et pourquoi elle pourrait impacter les mécanismes langagiers. DIVA/GODIVA permettent de situer le niveau de la production de parole où cette ressource peut exercer un effet.

Dans le cadre PRISM, l'amorce rythmique agit d'abord comme un stimulus temporel structuré. Elle peut faciliter l'extraction de régularités, soutenir la formation de prédictions temporelles et engager un couplage perception–action (Fiveash et al., 2021). Les résultats du chapitre mené auprès de tout-venants s'inscrivent dans cette logique : une structure rythmique perçue avant la parole peut influencer une action vocale ultérieure, et cette

influence semble pouvoir être associée, au moins en partie, à une compétence de synchronisation motrice. Même si les analyses EEG recueillies dans cette étude ne sont pas encore présentées dans le manuscrit, le cadre PRISM reste utile pour interpréter les résultats comportementaux. Il permet de comprendre l'amorçage comme une mise en état temporelle du système perceptivo-moteur, dont les marqueurs neurophysiologiques pourront être examinés ultérieurement à partir des données déjà acquises.

Les modèles DIVA/GODIVA permettent ensuite d'interpréter pourquoi l'effet comportemental apparaît surtout sur l'initiation. Dans ces modèles, la production de la parole implique la préparation de séquences, l'activation de programmes moteurs, leur déclenchement, puis leur ajustement sensorimoteur au cours de l'exécution (Bohland et al., 2010 ; Guenther, 2016 ; Tourville & Guenther, 2011). Les ganglions de la base y contribuent notamment à la sélection et au déclenchement des programmes moteurs. Dans la MP, ces mécanismes peuvent être fragilisés, ce qui rend particulièrement pertinent l'effet d'un repère temporel externe sur l'entrée en parole (Manes et al., 2024).

L'articulation des deux cadres conduit donc à la proposition suivante : l'amorçage rythmique ne modifie pas directement et uniformément la parole produite ; il modifie l'état temporel préparatoire du système perceptivo-moteur. Cet état préparatoire peut ensuite faciliter le déclenchement de la séquence parlée, surtout lorsque l'amorce fournit des repères suffisamment réguliers, saillants et exploitables. Les effets plus fins sur la prosodie, le débit ou les pauses dépendent ensuite de contraintes supplémentaires et ne se généralisent pas nécessairement.

Cette proposition répond à l'incertitude formulée dans l'introduction générale : aucun cadre théorique ne permettait à lui seul de rendre compte précisément de l'effet d'un rythme externe sur la production de la parole. Les résultats de la thèse montrent que PRISM et DIVA/GODIVA peuvent être mobilisés de manière complémentaire. PRISM éclaire le traitement du rythme ; DIVA/GODIVA éclairent le niveau de production où l'effet devient observable. Ensemble, ils permettent de dépasser une formulation générale du type « le rythme aide la parole » pour proposer une lecture plus précise : le rythme régulier peut préparer temporellement le système, et cette préparation se manifeste surtout au moment de l'initiation de la parole.

Cette articulation ouvre aussi une perspective mécanistique importante. Si le manuscrit montre principalement des effets comportementaux, l'étude menée auprès de tout-venants a également comporté un enregistrement EEG, dont les analyses ne sont pas présentées ici. Ces données permettront d'examiner ultérieurement le suivi neurophysiologique des régularités temporelles contenues dans les amorces. Plusieurs travaux suggèrent en effet que l'activité cérébrale peut s'aligner sur la structure temporelle d'un stimulus auditif et que cet alignement peut contribuer à la formation d'attentes temporelles pendant le traitement de la parole ou de séquences rythmiques (Kösem et al., 2018 ; L'Hermite & Zoefel, 2023 ; Zoefel et al., 2018). Dans cette perspective, des mesures comme la cohérence cerveau–stimulus, c'est-à-dire le degré de correspondance entre les variations de l'activité EEG et celles du signal auditif à certaines fréquences, pourraient permettre de tester plus directement certaines hypothèses issues du cadre PRISM, notamment le lien entre traitement auditif temporel, synchronisation aux régularités rythmiques et couplage sensorimoteur (Fiveash et al., 2021). Elles pourraient notamment aider à déterminer si les participants dont l'activité cérébrale suit davantage la structure temporelle des amorces sont aussi ceux qui présentent de meilleures compétences sensorimotrices ou une facilitation plus marquée de l'initiation de la parole. Dans la présente thèse, ces mécanismes sont donc mobilisés comme cadre interprétatif et comme perspective analytique ouverte par les données déjà recueillies, plutôt que comme résultat central.

Portée clinique

La thèse interrogeait aussi l'intérêt clinique des approches basées sur le rythme et leur place possible dans la pratique orthophonique. Les résultats expérimentaux ne permettent pas de proposer l'amorçage rythmique comme méthode clinique standardisée. Des effets globaux sur l'intelligibilité, l'expressivité ou la communication fonctionnelle n'ont pas été explorés. En revanche, ils permettent d'identifier une cible plausible : la préparation temporelle et l'initiation de la parole.

Dans la MP, les approches orthophoniques les mieux documentées restent centrées sur la voix, notamment LSVT LOUD. Les données disponibles, issues à la fois de revues de la littérature et d'essais randomisés récents, soutiennent l'efficacité de cette approche sur l'intensité vocale et sur l'impact fonctionnel des troubles vocaux (Atkinson-Clement et al.,

2015 ; Sackley et al., 2024). Ces approches restent centrales et les résultats de cette thèse ne les remettent pas en question. Ils suggèrent plutôt qu'un soutien rythmique pourrait constituer un complément ciblé, notamment lorsque les difficultés concernent l'initiation de la parole ou certaines dimensions prosodiques.

L'enquête menée auprès des orthophonistes apporte ici un éclairage complémentaire essentiel. Elle montre que la musique et le rythme sont déjà présents dans certaines pratiques cliniques, y compris auprès de patients présentant une dysarthrie, mais que ces usages restent hétérogènes, souvent peu formalisés, et limités par un manque de formation, d'outils stabilisés et de données sur l'efficacité. Ce résultat est important, car il situe les données expérimentales de la thèse dans un contexte clinique réel : le rythme n'apparaît pas comme une piste extérieure à la pratique orthophonique, mais comme une ressource déjà mobilisée, dont les modalités d'utilisation restent toutefois à préciser. Les effets observés dans les études expérimentales, en particulier sur l'initiation de la parole, peuvent donc être lus comme une première étape vers une meilleure formalisation de ces pratiques, plutôt que comme la proposition immédiate d'un protocole de rééducation.

Dans une perspective de pratique fondée sur les preuves, les résultats obtenus dans cette thèse invitent donc à une position prudente. Les résultats expérimentaux justifient de considérer le rythme comme une piste pertinente pour soutenir certaines dimensions temporelles de la parole, mais ils ne permettent pas d'en faire une recommandation clinique générale. L'intérêt clinique du rythme dépendra de sa capacité à être associé à des objectifs précis, à des profils de patients identifiés et à des critères d'évaluation fonctionnels.

Ainsi, le prolongement clinique de ce travail ne consiste pas à proposer directement un protocole d'intervention, mais à mieux définir les conditions dans lesquelles un soutien rythmique pourrait être pertinent. Il s'agit de passer d'une question générale : le rythme améliore-t-il la parole ? à une question plus précise : quels types de rythmes, pour quels profils de patients, sur quelles dimensions de parole, et avec quel impact perceptif ou fonctionnel ? La transition vers la clinique suppose de mieux caractériser les amorces, d'identifier les profils de patients les plus susceptibles d'en bénéficier, d'intégrer des mesures perceptives et fonctionnelles, et de mieux articuler les données expérimentales avec les usages réels des orthophonistes. C'est précisément sur ces conditions d'interprétation, de méthode et de

transfert que portent les limites générales et réflexions méthodologiques développées dans le chapitre suivant.

Chapitre 2
Limites générales et réflexions méthodologiques

Clarifier ce qui est manipulé dans les amorces rythmiques

Une première limite tient à la manière dont les amorces rythmiques sont définies et opérationnalisées. Dans les travaux sur l'amorçage rythmique, les notions de *rythme*, de *régularité* ou d'*irrégularité* sont souvent utilisées comme des catégories expérimentales relativement générales, alors qu'ils peuvent renvoyer à des propriétés acoustiques et temporelles distinctes. Deux amorces qualifiées de régulières peuvent différer par leur tempo, leur durée, la stabilité de leurs intervalles, la présence d'une pulsation saillante, leur organisation accentuelle ou encore leur proximité avec la structure prosodique de la phrase cible. De même, une amorce irrégulière peut correspondre à une absence de périodicité, à une perturbation locale d'un patron métrique, ou à une séquence qui conserve malgré tout certaines régularités exploitables par les participants.

Cette question est particulièrement importante lorsqu'on cherche à relier une amorce auditive à la production de la parole. Le rythme de la parole ne se réduit pas à une isochronie stricte des syllabes ou des accents, mais repose sur une organisation temporelle distribuée entre plusieurs niveaux : segments, syllabes, mots, groupes prosodiques, pauses et frontières prosodiques (Turk & Shattuck-Hufnagel, 2020). Les métriques acoustiques utilisées pour décrire cette organisation dépendent elles-mêmes fortement des choix de segmentation, des propriétés phonologiques de la langue, de la complexité syllabique et de la variabilité entre locuteurs (Arvaniti, 2012). Dans ce contexte, l'enjeu n'est pas seulement de comparer des conditions régulières et irrégulières, mais de préciser quelles propriétés temporelles de l'amorce sont susceptibles d'être encodées, anticipées, puis éventuellement transférées à la production.

Une conséquence méthodologique directe est de ne pas considérer l'opposition régulier/irrégulier comme suffisamment informative en elle-même. Une amorce irrégulière quasi aléatoire, une amorce métriquement structurée mais localement perturbée, ou une

séquence comportant des accents partiellement périodiques ne constituent pas des conditions équivalentes. Elles ne sollicitent pas les mêmes mécanismes perceptifs, n'imposent pas la même charge attentionnelle, et ne donnent pas aux participants les mêmes possibilités d'extraire un cadre temporel. Cette distinction est importante, car plusieurs travaux récents suggèrent que les effets de l'amorçage rythmique dépendent de propriétés fines des stimuli, notamment leur régularité, leur métricité, leur durée, leur structure hiérarchique et leur congruence avec la tâche linguistique ou verbale qui suit (Fiveash et al., 2020 ; Fiveash et al., 2021 ; György et al., 2024).

Dans les études futures, il serait donc utile de décrire plus systématiquement les amorces à partir de paramètres objectifs : distribution des intervalles, périodicité des accents, stabilité du beat, complexité métrique, durée totale, position du dernier événement auditif, richesse acoustique et degré de similarité avec la structure temporelle de la phrase cible. Une telle description permettrait de mieux interpréter les effets observés, en évitant d'attribuer à la seule « régularité » des effets qui pourraient dépendre plus précisément de la saillance du beat, de la durée d'exposition ou de la congruence entre amorce et production.

Plusieurs paramètres apparaissent particulièrement importants. Le tempo constitue une première variable clé. Les modèles distinguant le traitement d'une pulsation régulière et le traitement de durées absolues suggèrent que les réseaux impliqués peuvent varier selon que la tâche engage davantage une organisation *beat-based* ou *duration-based* (Teki et al., 2011, 2012). Dans le champ de la parole, cette distinction est centrale, car la fréquence pertinente n'est pas nécessairement la même selon que l'on cherche à soutenir les différents niveaux de la hiérarchie prosodique.

La durée de l'amorce constitue une deuxième variable majeure. Une amorce courte peut agir principalement comme un déclencheur temporel immédiat, alors qu'une amorce plus longue peut favoriser un état de préparation plus durable, une familiarisation avec la structure métrique, ou au contraire induire une fatigue ou une baisse attentionnelle. Cette question est importante pour interpréter les effets observés dans la thèse, mais aussi pour envisager leur prolongement. Les paradigmes utilisant des amorces brèves et ceux reposant sur des stimulations plus longues ne testent probablement pas exactement le même niveau de préparation temporelle. Les travaux sur l'amorçage rythmique en syntaxe, notamment

ceux issus de la littérature sur le *rhythmic priming effect*, invitent d'ailleurs à considérer la durée, la structure interne et la stabilité métrique des amorces comme des paramètres théoriques à part entière, et non comme de simples choix de présentation (Przybylski et al., 2013 ; Fiveash et al., 2020 ; György et al., 2024).

Une troisième dimension concerne la nature même des amorces. Une amorce purement rythmique, une amorce musicale, une amorce verbale ou une amorce composée de syllabes rythmiquement organisées n'engagent pas nécessairement les mêmes processus. Des travaux récents en population neurotypique suggèrent que la nature de l'amorce peut moduler le temps d'initiation de la parole, les amorces syllabiques semblant parfois offrir un compromis intéressant entre structuration temporelle et proximité avec la parole, sans introduire de contenu lexical ou syntaxique concurrent (Gindre, 2024). Ces résultats ne peuvent toutefois pas être directement transposés à la MP. Ils constituent plutôt un point de départ pour tester si, chez des patients présentant une atteinte des circuits impliqués dans la mise en route et la régulation temporelle de l'action, certaines amorces syllabiques ou non verbales sont plus facilement exploitables que des amorces musicales ou verbales complètes. Les amorces verbales apportent simultanément des informations prosodiques, phonologiques et parfois sémantiques. Elles peuvent donc être plus proches de la parole naturelle, mais plus difficiles à interpréter sur le plan mécanistique. À l'inverse, des amorces non verbales permettent de mieux isoler la dimension temporelle, mais peuvent être plus éloignées des situations cliniques ordinaires. Cette tension entre contrôle expérimental et validité écologique constitue un enjeu central pour les travaux futurs.

Le choix de la condition de comparaison constitue donc un enjeu distinct de la seule description des amorces. Une amorce irrégulière ne peut pas être considérée d'emblée comme une condition neutre : selon sa construction, elle peut réduire la prédictibilité temporelle, augmenter la charge attentionnelle, perturber la préparation de la réponse, ou au contraire conserver certaines régularités locales exploitables par les participants. Dès lors, un contraste entre amorce régulière et amorce irrégulière ne permet pas toujours de déterminer si l'effet observé reflète un bénéfice propre de la régularité, un coût lié à l'irrégularité, ou une combinaison de ces deux dimensions. L'ajout d'une condition silencieuse ou d'une stimulation auditive non rythmique, comme cela a été fait dans plusieurs travaux sur l'amorçage rythmique, permettrait de mieux distinguer l'effet spécifique de la structure

temporelle d'un effet plus général de stimulation auditive ou d'orientation attentionnelle. Cette logique a été adoptée dans certaines études d'amorçage rythmique, avec des conditions de silence (Ladányi et al., 2021 ; György et al., 2024), des scènes sonores environnementales sans régularité temporelle (Bedoin et al., 2016) ou des séquences texturales utilisées comme condition de comparaison auditive (Canette et al., 2020b).

Mieux articuler tâches contrôlées, mesures acoustiques et portée fonctionnelle

Une deuxième limite transversale concerne le choix des tâches de parole. Les études expérimentales de cette thèse reposent principalement sur des tâches contrôlées de lecture à voix haute. Ce choix présente des avantages méthodologiques évidents. Il permet de contrôler le contenu linguistique, de limiter la variabilité lexicale et syntaxique, de comparer les productions entre participants, et d'extraire des mesures acoustiques précises sur des phrases identiques, dont la structure pouvait être mise en relation avec celle des amorces perçues. Pour une question mécanistique portant sur l'initiation, le débit ou certains paramètres temporels, ce type de tâche est donc particulièrement pertinent.

Cependant, ce gain de contrôle s'accompagne d'un coût fonctionnel. La lecture à voix haute ne mobilise pas l'ensemble des contraintes qui structurent la communication ordinaire : choix lexical, planification du message, adaptation à l'interlocuteur, gestion des tours de parole, intention pragmatique, spontanéité et ajustement en temps réel. Les effets observés doivent donc être compris comme des effets obtenus dans une configuration expérimentale contrôlée, et non comme l'équivalent direct d'une amélioration de la communication quotidienne. Une réduction du temps d'initiation en lecture indique qu'une amorce rythmique peut faciliter le passage à la production dans cette tâche précise ; elle ne signifie pas nécessairement que le même effet se retrouverait dans une parole spontanée ou interactionnelle. Les travaux sur la parole dans la MP montrent en effet que les performances peuvent varier selon la tâche utilisée, par exemple entre lecture, monologue ou parole plus spontanée (Tjaden & Wilding, 2011 ; Barnish et al., 2017).

À cette question des tâches s'ajoute celle des critères d'évaluation. Une limite importante des études expérimentales présentées dans cette thèse réside dans l'écart possible entre variation acoustique, perceptibilité et signification clinique. Cette distinction est essentielle dans une perspective translationnelle. Dans la MP, les critères cliniquement

importants ne se réduisent pas à des paramètres acoustiques ou temporels isolés ; ils incluent aussi l'intelligibilité, l'effort de parole, l'effort d'écoute, la participation sociale et l'impact ressenti dans la vie quotidienne (Miller et al., 2006 ; Pu et al., 2021). Ces résultats apportent donc des éléments utiles pour comprendre les mécanismes en jeu, mais ils ne suffisent pas, à eux seuls, à conclure à un bénéfice fonctionnel direct.

Un enjeu important pour les travaux futurs sera donc d'articuler plus systématiquement les mesures acoustiques avec des mesures perceptives et fonctionnelles. De futures études pourraient intégrer des jugements d'auditeurs, par exemple sous la forme de tâches de transcription orthographique, d'évaluations de l'intelligibilité ou de la naturalité, ainsi que des auto-évaluations réalisées par les participants eux-mêmes. Les tâches de transcription orthographique constituent une approche fréquemment utilisée pour objectiver l'intelligibilité dans les troubles moteurs de la parole, car elles permettent de relier les caractéristiques des productions à ce qu'un auditeur parvient effectivement à comprendre (Chiu & Neel, 2020 ; Manes et al., 2026). Une telle triangulation permettrait de déterminer à partir de quel niveau une modification acoustique devient perceptible, puis éventuellement cliniquement significative.

Cette articulation est d'autant plus importante que des outils d'évaluation disponibles en français permettent aujourd'hui de mieux combiner mesures perceptives et acoustiques. Le protocole MonPaGe, par exemple, a été développé pour l'évaluation des troubles moteurs de la parole chez l'adulte francophone, en associant des tâches de parole standardisées à des mesures objectivables de différents paramètres, notamment prosodiques, articulatoires et temporels (Pernon et al., 2020 ; Laganaro et al., 2021). Dans de futures études, l'ajout d'un tel outil permettrait de mieux situer les effets expérimentaux observés dans un cadre d'évaluation plus proche des usages cliniques.

Chercher à préciser la variabilité interindividuelle

Une troisième limite concerne la caractérisation de la variabilité interindividuelle. Les résultats de la thèse montrent que l'amorçage rythmique ne produit pas un effet uniforme, ni entre les groupes, ni entre les participants d'un même groupe. Cette variabilité invite à dépasser une lecture moyenne des effets pour interroger les profils susceptibles de bénéficier

d’un repère temporel externe. Dans le cas de la MP, cette question est particulièrement importante, car l’hétérogénéité clinique peut concerner à la fois la sévérité motrice, le profil dominant des symptômes, les troubles cognitifs, les capacités auditives, la fatigue, les troubles de la parole et les compétences de synchronisation sensorimotrice.

Les travaux sur la synchronisation sensorimotrice montrent que la capacité à s’aligner sur un rythme externe varie fortement selon les individus (Repp, 2005 ; Repp & Su, 2013). Dans la MP, cette question a surtout été étudiée dans le champ du mouvement et de la marche, où les bénéfices de la SAR semblent en partie liés aux capacités individuelles de perception et de production du rythme (Dalla Bella et al., 2017 ; Cochen De Cock et al., 2018 ; Dalla Bella et al., 2018). De manière complémentaire, mais en dehors du champ de la MP, les données sur la synchronisation auditivo-vocale suggèrent l’existence de profils distincts dans la manière dont les locuteurs exploitent le rythme de la parole entendue pour guider leur propre production (Assaneo et al., 2019 ; Lizcano-Cortés et al., 2022). Dans le cadre de cette thèse, ces éléments conduisent à considérer les compétences rythmiques non comme de simples variables secondaires, mais comme des facteurs potentiels de modulation de l’effet d’amorçage. L’enjeu n’est donc pas seulement de savoir si une amorce régulière facilite la parole en moyenne, mais de déterminer chez quels participants, dans quelles conditions, et pour quelles dimensions de la production cette facilitation apparaît.

Dans les travaux futurs, il paraît nécessaire de formuler plus explicitement des hypothèses sur les profils de participants susceptibles de bénéficier d’un amorçage rythmique. Cette recommandation est particulièrement importante dans la MP, où l’hétérogénéité clinique est forte : sévérité motrice, profil dominant, troubles cognitifs, troubles auditifs et perceptifs, fatigue, médication, troubles axiaux ou troubles de la parole peuvent moduler la capacité à utiliser un support temporel externe.

Une limite des études présentées dans ce manuscrit est que ces dimensions individuelles n’ont pas toutes été documentées avec le même niveau de précision. Les questionnaires d’expérience musicale ou de sensibilité musicale apportent des informations utiles, mais ils reposent en partie sur des déclarations subjectives et ne permettent pas, à eux seuls, de caractériser les mécanismes perceptifs et sensorimoteurs impliqués dans le traitement du rythme. À l’inverse, les tâches de tapping fournissent des indices plus objectifs,

mais elles associent perception du rythme, programmation motrice et synchronisation gestuelle, ce qui rend parfois difficile l'identification du processus principalement impliqué.

Les travaux futurs gagneraient donc à combiner plusieurs niveaux d'évaluation. Des outils comme BAASTA permettent de caractériser les habiletés de perception et de synchronisation au beat, y compris dans des populations cliniques (Dalla Bella et al., 2017 ; Dalla Bella et al., 2024). Les tâches de synchronisation parole-parole peuvent documenter plus spécifiquement le couplage auditivo-vocal (Assaneo et al., 2019 ; Lizcano-Cortés et al., 2022). Des batteries perceptives musicales telles que PROMS ou ses versions modulaires permettent, quant à elles, d'évaluer plus directement les compétences auditives liées au rythme, au tempo ou à la hauteur (Law & Zentner, 2012 ; Strauss et al., 2024). L'association de ces mesures avec des variables cliniques (audiométrie, MoCA, UPDRS-III, profil moteur et sévérité de la dysarthrie) permettrait de mieux déterminer dans quelle mesure la sensibilité à l'amorçage dépend des propriétés de l'amorce, des compétences perceptives, rythmiques et sensorimotrices des participants, mais aussi de facteurs cliniques généraux susceptibles de moduler l'exploitation d'un stimulus temporel externe.

Consolider l'interprétation mécanistique

Une quatrième limite concerne l'interprétation mécanistique des effets observés. Les résultats présentés dans ce manuscrit soutiennent l'idée que l'amorçage rythmique agit principalement sur l'état préparatoire du système, avec un effet particulièrement visible sur le temps d'initiation de la parole. Toutefois, ils ne permettent pas encore de préciser entièrement par quels mécanismes ce soutien temporel se met en place, ni comment il se transforme en modulation comportementale.

Cette prudence est importante, car plusieurs mécanismes peuvent être envisagés. Une amorce rythmique pourrait agir par l'amélioration du traitement auditif précis, par la génération de prédictions temporelles, par l'entraînement d'oscillations neuronales, ou par le couplage sensorimoteur entre perception et action. Les résultats comportementaux et acoustiques présentés dans cette thèse ne permettent pas de dissocier directement ces mécanismes. Une réduction du temps d'initiation après amorce régulière est compatible avec une hypothèse de préparation temporelle, mais elle ne permet pas, à elle seule, de déterminer

si l'effet relève principalement de l'attention temporelle, de la prédiction, d'une facilitation motrice, ou d'une interaction entre ces différents niveaux.

L'articulation de ces mécanismes avec les modèles de production de la parole permet de situer plus précisément le niveau probable de l'effet dans la chaîne de production de la parole. Ces modèles suggèrent que la production parlée implique la préparation de séquences, l'activation de programmes moteurs, leur déclenchement, leur maintien et leur ajustement sensorimoteur. Dans la MP, l'atteinte des circuits impliquant les ganglions de la base peut fragiliser plusieurs de ces opérations, notamment l'initiation, le gating et le maintien du programme moteur en cours (Manes et al., 2024). Le fait que l'effet le plus robuste de l'amorçage porte sur le temps d'initiation s'accorde donc avec l'idée d'une action préférentielle sur la préparation et le déclenchement de la séquence parlée.

La comparaison entre les personnes atteintes de la MP et les personnes atteintes d'ASC apporte néanmoins un premier éclairage mécanistique. Elle montre que l'amorçage régulier facilite l'initiation de la parole dans les deux groupes cliniques, mais que les effets associés ne se limitent pas aux mêmes dimensions de la production. Dans la MP, le bénéfice sur l'initiation reste moins marqué que chez les participants contrôles, ce qui peut être mis en relation avec la fragilité des circuits cortico-striato-thalamo-corticaux impliqués dans la sélection et le déclenchement des programmes moteurs. Dans l'ASC, l'effet sur l'initiation est également présent, mais il s'accompagne d'effets sur certaines dimensions temporelles de l'énoncé, notamment le débit et les métriques syllabiques. Ce profil est compatible avec le rôle attribué au cervelet dans la précision temporelle, l'ajustement sensorimoteur et la coordination séquentielle (Breska & Ivry, 2018 ; Teki et al., 2011, 2012). Sans attribuer chaque effet à une structure cérébrale isolée, cette comparaison suggère que les atteintes des ganglions de la base et du cervelet modifient différemment la manière dont un contexte rythmique est utilisé pour préparer et organiser la parole.

Cependant, les données actuelles ne permettent pas encore d'établir une chaîne causale complète allant de la perception du rythme à la facilitation de la parole. Pour cela, il sera nécessaire d'intégrer dans les travaux futurs des mesures permettant de relier plus directement les propriétés objectives des amorces, les compétences rythmiques individuelles, les indices neurophysiologiques éventuels et les performances de parole. Dans cette perspective, des mesures de cohérence cerveau–stimulus ou d'entraînement neuronal

pourraient constituer un prolongement particulièrement pertinent, non comme une fin en soi, mais comme un moyen de tester l'hypothèse qu'une structure rythmique externe peut induire un état temporel préparatoire susceptible de persister brièvement et de moduler la tâche suivante.

Cette perspective devra toutefois être abordée avec prudence. Les indices neurophysiologiques de cohérence cerveau–stimulus sont sensibles au choix des fréquences d'intérêt, à la qualité du signal EEG, à la durée des stimuli et aux propriétés acoustiques des amorces. De plus, un marqueur de synchronisation cerveau–stimulus ne permet pas, à lui seul, de déterminer le mécanisme exact sous-jacent. Il peut refléter le suivi acoustique du stimulus, l'alignement de l'activité cérébrale sur sa structure temporelle, ou encore des processus attentionnels et prédictifs liés à la régularité perçue. Il ne permet donc pas nécessairement de conclure à un entraînement neuronal actif, ni d'établir directement que ce suivi neurophysiologique est responsable de l'effet comportemental observé sur la parole. Les travaux sur l'entraînement neuronal soulignent d'ailleurs que ces indices doivent être interprétés avec prudence, car ils peuvent participer aux mécanismes de sélection temporelle de l'information sans constituer, à eux seuls, une preuve directe de facilitation sensorimotrice ou comportementale (Calderone et al., 2014).

La recommandation principale est donc de renforcer cette mise en regard entre indices neurophysiologiques, compétences rythmiques et production de la parole. Les futurs travaux devraient chercher à relier explicitement les propriétés des amorces, les mesures de perception et de synchronisation rythmiques, les marqueurs neurophysiologiques éventuels et les effets comportementaux sur la parole. L'objectif ne serait pas seulement de montrer que ces niveaux covarient, mais de déterminer s'ils permettent de prédire la sensibilité comportementale aux amorces. Dans cette logique, il sera important de distinguer plusieurs rôles possibles des marqueurs neurophysiologiques : ils peuvent être de simples corrélats du traitement rythmique, des modérateurs de l'effet comportemental, ou des médiateurs plausibles entre amorce et production. Ces possibilités ne sont pas équivalentes et doivent être testées par des modèles explicitement construits pour cela.

Ainsi, les résultats actuels soutiennent une lecture selon laquelle le rythme agit principalement sur l'état préparatoire du système, mais ils ne permettent pas encore de conclure à un mécanisme unique. Cette limite ne fragilise pas l'interprétation générale ; elle

précise au contraire les étapes nécessaires pour passer d'une démonstration comportementale à une compréhension mécanistique plus complète.

Encadrer la portée clinique

Une cinquième limite concerne la transposition clinique. Les résultats de la thèse montrent que l'amorçage rythmique peut influencer certaines dimensions de la parole, mais ils ne permettent pas encore de définir une intervention orthophonique stabilisée. Cette distinction est essentielle. Une démonstration expérimentale, même cohérente, ne suffit pas à fonder une recommandation clinique. Pour qu'un paradigme puisse être envisagé dans une perspective de rééducation, plusieurs conditions supplémentaires doivent être réunies : identification des patients susceptibles de répondre, définition des objectifs thérapeutiques, choix du type de stimulation, tolérance et acceptabilité, durée d'utilisation réaliste, critères d'efficacité fonctionnels, et comparaison avec les approches déjà disponibles.

À ce stade, la portée clinique la plus raisonnable des résultats concerne la facilitation de l'entrée en parole. L'effet le plus robuste observé dans la thèse porte sur le temps d'initiation, tandis que les effets sur le débit, les métriques rythmiques ou les dimensions prosodiques plus fines sont plus conditionnels. Cela suggère que l'amorçage rythmique ne doit pas être présenté comme une méthode générale d'amélioration de la parole ou de la prosodie dans la MP. Il peut plutôt être envisagé comme un outil potentiel de préparation temporelle, susceptible de soutenir certaines dimensions spécifiques de la production, chez certains profils de participants, dans certaines configurations de stimulation.

Cette position est importante dans une perspective de pratique fondée sur les preuves. Les interventions orthophoniques doivent s'appuyer non seulement sur des résultats expérimentaux, mais aussi sur leur pertinence fonctionnelle, leur faisabilité et leur acceptabilité. Dans la MP, les approches les mieux documentées restent aujourd'hui centrées sur la voix et l'intensité vocale, notamment LSVT LOUD, avec des effets rapportés sur l'intensité, certaines mesures de parole et le retentissement fonctionnel perçu (Ramig et al., 2001 ; Pu et al., 2021 ; Sackley et al., 2024). Les résultats de cette thèse ne remettent pas en cause ces approches ; ils suggèrent plutôt qu'un soutien rythmique pourrait, à terme, constituer un complément ciblé pour certaines dimensions temporelles de la parole.

L'enquête menée auprès des orthophonistes montre par ailleurs que la musique et le rythme sont déjà présents dans certaines pratiques, mais de manière hétérogène, avec des objectifs, des modalités et des niveaux de formalisation variables. Cette situation rejoint les constats plus généraux formulés dans les revues récentes sur les interventions utilisant des éléments musicaux ou rythmiques en orthophonie : ces approches sont prometteuses, mais restent souvent difficiles à comparer en raison de la diversité des protocoles, des populations et des critères d'évaluation (Ong et al., 2024). Ces éléments suggèrent un intérêt clinique réel, mais aussi la nécessité de mieux encadrer les usages.

La recommandation qui en découle est de ne pas opposer recherche fondamentale et clinique, mais de construire des étapes intermédiaires. Avant d'envisager une intervention, il faut déterminer quelles stimulations sont les plus pertinentes, pour quels profils de patients, avec quels effets mesurables, et selon quels critères perceptifs ou fonctionnels. Cela suppose d'intégrer dès les études expérimentales des mesures qui dépassent le seul signal acoustique : intelligibilité, jugement perceptif, ressenti du patient, effort de parole, effort d'écoute et impact subjectif.

Il est également nécessaire de penser la faisabilité clinique des stimulations proposées. Des amorces très contrôlées sur le plan expérimental ne sont pas toujours transposables telles quelles en séance. À l'inverse, des stimulations plus naturelles, verbales, musicales ou syllabiques, peuvent être plus faciles à intégrer en pratique, mais elles rendent plus complexe l'identification du mécanisme actif. Ce compromis entre contrôle expérimental et pertinence clinique devra être au cœur des travaux futurs. Il ne s'agira pas de rechercher un stimulus universel, mais d'identifier des configurations de stimulation adaptées à des objectifs précis et à des profils de patients particuliers.

Ainsi, les résultats de la thèse ouvrent une perspective clinique, mais ne la réalisent pas encore. Ils fournissent un cadre théorique et expérimental plus solide pour poser la question suivante : comment passer d'un effet d'amorçage observé en laboratoire à une utilisation ciblée, évaluée et pertinente dans le champ orthophonique ?

Chapitre 3
Perspectives

Ce chapitre final présente une perspective sous forme d'une étude expérimentale pour approfondir les résultats observés et examiner la pertinence clinique de l'amorçage rythmique. Comme discuté dans les chapitres précédents, plusieurs interrogations théoriques demeurent ouvertes dans le champ des stimulations auditives rythmiques : les mécanismes cognitifs et neuronaux qui sous-tendent leurs effets, la nature des paramètres rythmiques les plus efficaces, et les raisons pour lesquelles certains individus y répondent mieux que d'autres. Les deux études empiriques présentées dans ce manuscrit explorent chacune à sa façon ces éléments. Une perspective serait d'associer leurs méthodologies pour aller plus avant dans l'évaluation de l'effet de l'amorçage rythmique sur la production de la parole chez les personnes atteintes de la MP tout en caractérisant mieux les facteurs individuels susceptibles de moduler cet effet et à identifier des configurations de stimulation potentiellement pertinentes dans une perspective de recherches translationnelles ou cliniques ultérieures.

Sur le plan méthodologique, l'un des prolongements les plus directs de cette thèse serait de comparer, chez les mêmes participants, plusieurs formats d'amorces rythmiques dont les effets ont jusqu'ici été étudiés de manière séparée. Les résultats présentés dans ce manuscrit suggèrent que l'amorçage régulier influence surtout l'initiation de la parole, tandis que les effets sur le débit, les métriques rythmiques et les dimensions prosodiques restent plus conditionnels. Une étude future pourrait donc chercher à déterminer si cette facilitation dépend principalement de la régularité temporelle, de la durée d'exposition, de la nature du stimulus ou de sa proximité avec la parole.

Une première manipulation concernerait la durée des amorces. Les amorces brèves, proches de celles utilisées dans les études expérimentales de cette thèse, pourraient être comparées à des séquences plus longues, de l'ordre de trente secondes, telles qu'elles sont fréquemment utilisées dans les travaux sur l'amorçage rythmique et le traitement syntaxique (Przybylski et al., 2013 ; Fiveash et al., 2020 ; György et al., 2024). Cette comparaison permettrait de tester si une exposition courte suffit à faciliter le déclenchement de la parole,

ou si une exposition plus longue favorise un état de préparation temporelle plus durable. Elle permettrait également de distinguer deux niveaux potentiellement différents d'effet : un effet local, immédiatement lié à la transition entre l'amorce et la phrase, et un effet plus global, lié à l'installation progressive d'un cadre temporel ou métrique.

Au-delà de la durée des amorces, une deuxième manipulation pourrait concerner leur nature. À notre connaissance, la comparaison directe entre amorces instrumentales, amorces syllabiques non lexicales et amorces verbales complètes n'a pas encore été réalisée dans le cadre de la MP. Pourtant, plusieurs travaux suggèrent que les patients peuvent exploiter des stimulations rythmiques proches de la parole, notamment dans des paradigmes d'accommodation ou de complétion à partir d'un modèle auditif (Aichert et al., 2021 ; Späth et al., 2016, 2022). En population générale, d'autres résultats indiquent également que la nature de l'amorce peut moduler son effet sur le temps d'initiation de la parole, avec un bénéfice plus marqué pour certaines amorces vocales ou non langagières que pour des amorces purement musicales ou verbales (Gindre, 2024). Comparer directement ces différents formats permettrait donc de mieux distinguer ce qui relève de la structure temporelle elle-même de ce qui dépend de la proximité avec la parole.

Dans une perspective orthophonique, il ne s'agirait pas de reprendre directement les paradigmes d'Aichert et Späth à partir de stimuli langagiers complets. Si ces stimuli présentent une certaine validité écologique, ils rendent l'interprétation plus complexe, car leur structure rythmique se mêle à des informations prosodiques, phonologiques, sémantiques et syntaxiques. Une alternative consisterait à utiliser des séquences syllabiques rythmiquement organisées, proches de celles proposées par Gindre (2024). Ces stimuli conservent une dimension vocale et cliniquement mobilisable, tout en limitant l'influence du contenu lexical et sémantique. Ils pourraient ainsi constituer un compromis intéressant entre contrôle expérimental et transférabilité clinique.

L'ajout d'une composante motrice constitue une autre piste, mais elle devrait être envisagée avec prudence. Plusieurs travaux suggèrent que l'engagement moteur, par exemple sous la forme d'un tapping synchronisé ou d'un entraînement audio-moteur préalable, peut renforcer certains effets du rythme sur le traitement de la parole (Cason et al., 2015 ; Falk & Dalla Bella, 2016 ; Falk et al., 2017 ; Berthault et al., 2024) et sa production (Gindre, 2024). Toutefois, dans le contexte de la MP, une telle manipulation pourrait introduire une charge

motrice supplémentaire et interagir avec les symptômes moteurs des patients. Elle ne devrait donc pas être considérée uniquement comme un moyen d'augmenter l'efficacité de l'amorce, mais comme une condition expérimentale distincte, engageant potentiellement d'autres mécanismes que l'écoute passive. Dans une perspective clinique, il serait ainsi nécessaire de déterminer si l'ajout d'un geste synchronisé facilite réellement la préparation de la parole, ou s'il augmente au contraire la complexité de la tâche pour certains patients.

Concernant l'évaluation initiale des participants, deux éléments pourraient être renforcés dans les travaux futurs. Le premier concerne l'évaluation de la parole avant toute stimulation rythmique. Dans les études présentées ici, une évaluation clinique de la parole était déjà réalisée, notamment à l'aide de la BECD (Auzou & Rolland-Monnoury, 2019). Toutefois, il pourrait être utile d'ajouter un bloc de lecture en silence, construit sur le même format que les blocs expérimentaux avec amorces, mais présenté avant toute exposition aux stimuli rythmiques. Une telle mesure permettrait de disposer d'un niveau de référence plus directement comparable aux conditions expérimentales, tout en limitant le risque qu'un effet de persistance lié aux amorces précédentes influence la condition sans stimulation.

Le second élément concerne la caractérisation clinique de la parole. L'utilisation d'un outil standardisé comme MonPaGe pourrait constituer un complément intéressant, dans la mesure où il propose une évaluation perceptive et acoustique structurée des troubles moteurs de la parole chez l'adulte francophone, avec des tâches standardisées et des mesures objectivables de plusieurs dimensions de la production (Laganaro et al., 2021 ; Pernon et al., 2020). Il ne faudrait toutefois pas en faire un outil central pour l'analyse fine de la prosodie. Dans l'étude de validation du protocole de screening, les auteurs indiquent que les contrastes prosodiques étaient déviants chez peu de patients, ce qui pourrait refléter une sensibilité limitée de ces mesures pour certains profils de troubles moteurs de la parole (Laganaro et al., 2021). MonPaGe serait donc particulièrement utile pour situer les participants dans un cadre clinique normé et limiter la part de jugement subjectif, mais il ne pourrait se substituer à des mesures acoustiques ciblées sur les paramètres prosodiques d'intérêt.

En complément des blocs silences, l'utilisation de stimuli non rythmiques, tel que du bruits rose ou des sons d'environnement, permettrait de faire la part des choses entre stimuli rythmiques, irréguliers comme réguliers, et stimulation auditive sans structure rythmique. Cet

ajout permettrait ainsi de bien différencier les aspects attentionnels liés à une stimulation, qu'importe sa nature, avec l'effet réel du rythme et permettrait d'affiner la différenciation régulier/irrégulier. Ces ajouts sont particulièrement importants, car ils répondent directement à l'une des limites des études antérieures : le fait de ne pouvoir tester qu'un nombre encore restreint de formats de stimulation, laissant ouverte la question pourtant essentielle de la forme optimale de l'amorce. Il faut préciser ici que l'objectif n'est pas de chercher le « stimulus universel », mais plutôt d'examiner une interaction potentielle entre propriétés des amorces et caractéristiques des participants.

S'agissant de ces dernières, l'étude réalisée auprès des patients n'avait pas proposé d'évaluation des compétences de synchronisation motrice. De plus, nous n'avons pas utilisé dans les deux études empiriques des outils d'évaluation « objectifs » des compétences musicales. Certes les tâches de tapping sont présentes ; mais elles contiennent une composante de production, et le BST de la H-BAT (Fujii & Schlaug, 2013) fournit des seuils de perception et de production du beat, ne renseignant pas directement sur les compétences musicales. Nous proposons donc de construire une évaluation de cet ensemble à partir de différents outils complémentaires. Les capacités de perception et de synchronisation au beat peuvent être documentées à l'aide de la version mobile abrégée de la BAASTA (Dalla Bella et al., 2024) qui propose le score de Beat Tracking Index. La tâche Speech-to-Speech Synchronization permettra de documenter le couplage auditivo-vocal (Lizcano-Cortés et al., 2022) et pourra être particulièrement mise en regard avec les stimuli rythmiques non langagiers. Les compétences perceptives musicales seront évaluées à l'aide de la Modular-PROMS (Strauss et al., 2024), batterie permettant notamment d'évaluer entre autres la perception du rythme mais aussi du tempo ou de mélodies. L'histoire musicale ainsi que la sensibilité à la récompense musicale seront documentées par un questionnaire de pratique musicale et par le BMRQ (Saliba et al., 2016). Des variables de contrôle devront compléter cette caractérisation : audiométrie, MoCA (Nasreddine et al., 2005), score UPDRS-III (Goetz et al., 2008). Ce score permet par ailleurs de définir précisément la symptomatologie motrice des patients et donc nous permettrait de créer des clusters bien que pour ce faire, un nombre conséquent d'inclusions devrait être envisagé.

Ce choix méthodologique est particulièrement cohérent avec la logique des travaux présentés dans cette thèse, car il permet de relier les effets observés sur la parole à des ressources perceptives, motrices, cognitives et musicales susceptibles d'en expliquer la variabilité. Avec le cadre expérimental proposé ici, les compétences rythmiques deviennent des composantes centrales de l'interprétation de l'effet d'amorçage rythmique dans la MP.

Dans une perspective clinique, l'évaluation des effets ne peut pas se limiter aux mesures acoustiques. Une modification du temps d'initiation, du débit ou d'un paramètre prosodique n'a de portée clinique que si elle s'accompagne d'un effet perceptible, fonctionnellement pertinent, ou ressenti comme utile par le patient. Cette distinction est importante, car certaines variations acoustiques peuvent être statistiquement détectables sans pour autant modifier l'intelligibilité, la naturalité de la parole ou l'effort de communication. À l'inverse, des effets acoustiques modestes peuvent devenir pertinents s'ils améliorent la fluidité perçue, la facilité d'entrée dans la parole ou le confort de communication.

En cohérence avec les limites discutées précédemment, cette étude devrait également intégrer des critères perceptifs et fonctionnels. Les mesures acoustiques seraient ainsi complétées par des tâches d'intelligibilité, des jugements de naturalité, d'effort d'écoute ou de fluidité, ainsi que par des questionnaires recueillant le ressenti des participants. L'objectif serait de déterminer si les effets observés sur le signal de parole restent de simples modifications acoustiques mesurables en laboratoire, ou s'ils peuvent préfigurer un intérêt clinique réel.

Une telle étude permettrait d'identifier les marqueurs pertinents, de documenter les modérateurs individuels et de comparer des modalités de stimulation susceptibles d'être utiles en rééducation. En ce sens, la perspective ouverte est celle d'un passage d'une démonstration de l'effet à une compréhension plus fine de ses conditions d'efficacité, préalable indispensable à toute proposition raisonnée d'intervention fondée sur le rythme chez les personnes atteintes de la MP.

Références

Ackermann, H., & Ziegler, W. (1994). Acoustic analysis of vocal instability in cerebellar dysfunctions. *Annals of Otology, Rhinology & Laryngology, 103*(2), 98–104. https://doi.org/10.1177/000348949410300203

Aichert, I., Lehner, K., Falk, S., Späth, M., Franke, M., & Ziegler, W. (2021). In time with the beat: Entrainment in patients with phonological impairment, apraxia of speech, and Parkinson's disease. *Brain Sciences, 11*(11), 1524. https://doi.org/10.3390/brainsci11111524

Altmann, L. J. P., & Troche, M. S. (2011). High-level language production in Parkinson's disease: A review. *Parkinson's Disease, 2011*, 238956. https://doi.org/10.4061/2011/238956

Armstrong, M. J., & Okun, M. S. (2020). Diagnosis and treatment of Parkinson disease: A review. *JAMA, 323*(6), 548–560. https://doi.org/10.1001/jama.2019.22360

Arvaniti, A. (2012). The usefulness of metrics in the quantification of speech rhythm. *Journal of Phonetics, 40*(3), 351–373. https://doi.org/10.1016/j.wocn.2012.02.003

Arvaniti, A. (2020). The phonetics of prosody. In M. Aronoff (Ed.), *Oxford Research Encyclopedia of Linguistics*. Oxford University Press. https://doi.org/10.1093/acrefore/9780199384655.013.411

Assaneo, M. F., Ripollés, P., Orpella, J., Lin, W. M., de Diego-Balaguer, R., & Poeppel, D. (2019). Spontaneous synchronization to speech reveals neural mechanisms facilitating language learning. *Nature Neuroscience, 22*(4), 627–632. https://doi.org/10.1038/s41593-019-0353-z

Astésano, C., & Bertrand, R. (2016). Accentuation et niveaux de constituance en français : Enjeux phonologiques et psycholinguistiques. *Langue française, 191*(3), 11–30. https://doi.org/10.3917/lf.191.0011

Atalar, M. S., Oguz, O., & Genc, G. (2023). Hypokinetic dysarthria in Parkinson's disease: A narrative review. *Medical Bulletin of Sisli Etfal Hospital, 57*(2), 163–170. https://doi.org/10.14744/SEMB.2023.29560

Atkinson-Clement, C., Sadat, J., & Pinto, S. (2015). Behavioral treatments for speech in Parkinson's disease: Meta-analyses and review of the literature. *Neurodegenerative Disease Management, 5*(3), 233–248. https://doi.org/10.2217/nmt.15.16

Atkinson-Clement, C., Letanneux, A., Baille, G., Cuartero, M.-C., Véron-Delor, L., Robieux, C., Berthelot, M., Robert, D., Azulay, J.-P., Defebvre, L., Ferreira, J., Eusebio, A., Moreau, C., & Pinto, S. (2019). Psychosocial impact of dysarthria: The patient-reported outcome as part of the clinical management. *Neurodegenerative Diseases, 19*(1), 12–21. https://doi.org/10.1159/000499627

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

Baez-Ortega, A. (2018, May 28). *Bayesian robust correlation with Stan in R*. In silico. https://baezortega.github.io/2018/05/28/robust-correlation/

Barnish, M. S., Whibley, D., Horton, S. M. C., Butterfint, Z. R., & Deane, K. H. O. (2016). Roles of cognitive status and intelligibility in everyday communication in people with Parkinson's disease: A systematic review. *Journal of Parkinson's Disease, 6*(2), 453–462. https://doi.org/10.3233/JPD-150757

Barnish, M. S., Horton, S. M. C., Butterfint, Z. R., Clark, A. B., Atkinson, R. A., & Deane, K. H. O. (2017). Speech and communication in Parkinson's disease: A cross-sectional exploratory study in the UK. *BMJ Open, 7*(5), e014642. https://doi.org/10.1136/bmjopen-2016-014642

Basirat, A. (2025). Language in Parkinson's disease. In L. Cummings (Ed.), *The Oxford handbook of communication disorders in neurodegenerative diseases.* Oxford University Press. https://doi.org/10.1093/oxfordhb/9780198888482.013.0013

Baudouin, R., Lechien, J. R., Carpentier, L., Gurruchaga, J.-M., Lisan, Q., & Hans, S. (2023). Deep brain stimulation impact on voice and speech quality in Parkinson's disease: A systematic review. *Otolaryngology–Head and Neck Surgery, 168*(3), 307–318. https://doi.org/10.1177/01945998221120189

Baylor, C., Cook, K. J., & McAuliffe, M. J. (2024). "Take us into account": Perspectives of family members of people with Parkinson's disease regarding speech-language pathology intervention. *American Journal of Speech-Language Pathology, 33*(2), 952–968. https://doi.org/10.1044/2023_AJSLP-23-00273

Bedoin, N., Brisseau, L., Molinier, P., Roch, D., & Tillmann, B. (2016). Temporally regular musical primes facilitate subsequent syntax processing in children with specific language impairment. *Frontiers in Neuroscience, 10*, 245. https://doi.org/10.3389/fnins.2016.00245

Bégel, V., Seilles, A., & Dalla Bella, S. (2018). Rhythm Workers: A music-based serious game for training rhythm skills. *Music & Science, 1*, 1–16. https://doi.org/10.1177/2059204318794369

Benedetti, V., Weill-Chounlamountry, A., Pradat-Diehl, P., & Villain, M. (2022). Assessment tools and rehabilitation treatments for aprosodia following acquired brain injury: A scoping review. *International Journal of Language & Communication Disorders, 57*(3), 474–496. https://doi.org/10.1111/1460-6984.12692

Benoit, C.-E., Dalla Bella, S., Farrugia, N., Obrig, H., Mainka, S., & Kotz, S. A. (2014). Musically cued gait-training improves both perceptual and motor timing in Parkinson's disease. *Frontiers in Human Neuroscience, 8*, 494. https://doi.org/10.3389/fnhum.2014.00494

Berthault, E., Chen, S., Falk, S., Morillon, B., & Schön, D. (2024). Auditory and motor priming of metric structure improves understanding of degraded speech. *Cognition, 248*, 105793. https://doi.org/10.1016/j.cognition.2024.105793

Bloem, B. R., Okun, M. S., & Klein, C. (2021). Parkinson's disease. *The Lancet, 397*(10291), 2284–2303. https://doi.org/10.1016/S0140-6736(21)00218-X

Bögels, S. , & Torreira, F. (2021). Turn-end estimation in conversational turn-taking: The roles of context and prosody. *Discourse Processes, 58*(10), 903–924. https://doi.org/10.1080/0163853X.2021.1986664

Bohland, J. W., Bullock, D., & Guenther, F. H. (2010). Neural representations and mechanisms for the performance of simple speech sequences. *Journal of Cognitive Neuroscience, 22*(7), 1504–1529. https://doi.org/10.1162/jocn.2009.21306

Bourdon, A., Damm, L., Dotov, D., Ihalainen, P., Dalla Bella, S., Bardy, B. G., & Cochen De Cock, V. (2025). Gait ecological assessment in persons with Parkinson's disease engaged in a synchronized musical rehabilitation program. *npj Parkinson's Disease, 11*, 12. https://doi.org/10.1038/s41531-024-00852-6

Brabenec, L., Mekyska, J., Galaz, Z., & Rektorova, I. (2017). Speech disorders in Parkinson's disease: Early diagnostics and effects of medication and brain stimulation. *Journal of Neural Transmission, 124*(3), 303–334. https://doi.org/10.1007/s00702-017-1676-0

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision, 10*(4), 433–436. https://doi.org/10.1163/156856897X00357

Breitenstein, C., Van Lancker, D., Daum, I., & Waters, C. H. (2001). Impaired perception of vocal emotions in Parkinson's disease: Influence of speech time processing and executive functioning. *Brain and Cognition, 45*(2), 277–314. https://doi.org/10.1006/brcg.2000.1246

Breska, A., & Ivry, R. B. (2018). Double dissociation of single-interval and rhythmic temporal prediction in cerebellar degeneration and Parkinson's disease. *Proceedings of the National Academy of Sciences of the United States of America, 115*(48), 12283–12288. https://doi.org/10.1073/pnas.1810596115

Brin-Henry, F., Courrier, C., Lederlé, E., & Masy, V. (2011). *Dictionnaire d'orthophonie* (4e éd.). Ortho Édition.

Bürkner, P.-C. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software, 80*(1), 1–28. https://doi.org/10.18637/jss.v080.i01

Van der Bruggen, S., De Letter, M., & Rietveld, T. (2024). Effects of near-monotonous speech of persons with Parkinson's disease on listening effort and intelligibility. *Clinical Linguistics & Phonetics, 38*(10), 935–948. https://doi.org/10.1080/02699206.2023.2272032

Calderone, D. J., Lakatos, P., Butler, P. D., & Castellanos, F. X. (2014). Entrainment of neural oscillations as a modifiable substrate of attention. *Trends in Cognitive Sciences, 18*(6), 300–309. https://doi.org/10.1016/j.tics.2014.02.005

Canette, L.-H., Bedoin, N., Lalitte, P., Bigand, E., & Tillmann, B. (2019). The regularity of rhythmic primes influences syntax processing in adults. *Auditory Perception & Cognition, 2*(3), 163–179. https://doi.org/10.1080/25742442.2020.1752080

Canette, L.-H., Fiveash, A., Krzonowski, J., Corneyllie, A., Lalitte, P., Thompson, D., Trainor, L., Bedoin, N., & Tillmann, B. (2020a). Regular rhythmic primes boost P600 in grammatical error processing in dyslexic adults and matched controls. *Neuropsychologia, 138*, 107324. https://doi.org/10.1016/j.neuropsychologia.2019.107324

Canette, L.-H., Lalitte, P., Bedoin, N., Pineau, M., Bigand, E., & Tillmann, B. (2020b). Rhythmic and textural musical sequences differently influence syntax and semantic processing in children. *Journal of Experimental Child Psychology, 191*, 104711. https://doi.org/10.1016/j.jecp.2019.104711

Cason, N., & Schön, D. (2012). Rhythmic priming enhances the phonological processing of speech. *Neuropsychologia, 50*(11), 2652–2658. https://doi.org/10.1016/j.neuropsychologia.2012.07.018

Cason, N., Astésano, C., & Schön, D. (2015a). Bridging music and speech rhythm: Rhythmic priming and audio–motor training affect speech perception. *Acta Psychologica, 155*, 43–50. https://doi.org/10.1016/j.actpsy.2014.12.002

Cason, N., Hidalgo, C., Isoard, F., Roman, S., & Schön, D. (2015b). Rhythmic priming enhances speech production abilities: Evidence from prelingually deaf children. *Neuropsychology, 29*(1), 102–107. https://doi.org/10.1037/neu0000115

Cheang, H. S., & Pell, M. D. (2007). An acoustic investigation of Parkinsonian speech in linguistic and emotional contexts. *Journal of Neurolinguistics, 20*(3), 221–241. https://doi.org/10.1016/j.jneuroling.2006.07.001

Chen, X., & Sidtis, D. (2024). Contrastive stress in persons with Parkinson's disease who speak Mandarin: Task effect in production and preserved perception. *Journal of Neurolinguistics, 69*, 101173. https://doi.org/10.1016/j.jneuroling.2023.101173

Chiu, Y.-F., & Neel, A. (2020). Predicting intelligibility deficits in Parkinson's disease with perceptual speech ratings. *Journal of Speech, Language, and Hearing Research, 63*(2), 433–443. https://doi.org/10.1044/2019_JSLHR-19-00134

Cloud, C., Georgen-Schwartz, K., & Hilger, A. (2024). The contributions of pitch, loudness, and rate control to speech naturalness in cerebellar ataxia. *American Journal of Speech-Language Pathology, 33*(5), 2536–2555. https://doi.org/10.1044/2024_AJSLP-24-00018

Cochen De Cock, V., Dotov, D. G., Ihalainen, P., Bégel, V., Galtier, F., Lebrun, C., Picot, M.-C., Driss, V., Landragin, N., Geny, C., Bardy, B. G., & Dalla Bella, S. (2018). Rhythmic abilities and musical training in Parkinson's disease: Do they help? *npj Parkinson's Disease, 4*, 8. https://doi.org/10.1038/s41531-018-0043-7

Cochen De Cock, V., Dotov, D., Damm, L., Lacombe, S., Ihalainen, P., Picot, M.-C., Galtier, F., Lebrun, C., Giordano, A., Driss, V., Geny, C., Garzo, A., Hernandez, E., Van Dyck, E., Leman, M., Villing, R., Bardy, B. G., & Dalla Bella, S. (2021). BeatWalk: Personalized music-based gait rehabilitation in Parkinson's disease. *Frontiers in Psychology, 12*, 655121. https://doi.org/10.3389/fpsyg.2021.655121

Coulombe, V., Joyal, M., Martel-Sauvageau, V., & Monetta, L. (2023). Affective prosody disorders in adults with neurological conditions: A scoping review. *International Journal of Language & Communication Disorders, 58*(6), 2063–2099. https://doi.org/10.1111/1460-6984.12909

Criscuolo, A., Schwartze, M., Nozaradan, S., & Kotz, S. A. (2025). Basal ganglia and cerebellar lesions causally impact the neural encoding of temporal regularities. *Imaging Neuroscience, 3*, 00492. https://doi.org/10.1162/imag_a_00492

Dalla Bella, S., Benoit, C.-E., Farrugia, N., Schwartze, M., & Kotz, S. A. (2015). Effects of musically cued gait training in Parkinson's disease: Beyond a motor benefit. *Annals of the New York Academy of Sciences, 1337*(1), 77–85. https://doi.org/10.1111/nyas.12651

Dalla Bella, S., Benoit, C.-E., Farrugia, N., Keller, P. E., Obrig, H., Mainka, S., & Kotz, S. A. (2017a). Gait improvement via rhythmic stimulation in Parkinson's disease is linked to rhythmic skills. *Scientific Reports, 7*, 42005. https://doi.org/10.1038/srep42005

Dalla Bella, S., Farrugia, N., Benoit, C.-E., Bégel, V., Verga, L., Harding, E., & Kotz, S. A. (2017b). BAASTA: Battery for the Assessment of Auditory Sensorimotor and Timing Abilities. *Behavior Research Methods, 49*(3), 1128–1145. https://doi.org/10.3758/s13428-016-0773-6

Dalla Bella, S., Dotov, D., Bardy, B., & Cochen de Cock, V. (2018). Individualization of music-based rhythmic auditory cueing in Parkinson's disease. *Annals of the New York Academy of Sciences, 1423*(1), 308–317. https://doi.org/10.1111/nyas.13859

Dalla Bella, S., Foster, N. E. V., Laflamme, H., Zagala, A., Melissa, K., Komeilipoor, N., Blais, M., Rigoulot, S., & Kotz, S. A. (2024). Mobile version of the Battery for the Assessment of Auditory Sensorimotor and Timing Abilities (BAASTA): Implementation and adult norms. *Behavior Research Methods, 56*(4), 3737–3756. https://doi.org/10.3758/s13428-024-02363-x

Darley, F. L., Aronson, A. E., & Brown, J. R. (1969a). Clusters of deviant speech dimensions in the dysarthrias. *Journal of Speech and Hearing Research, 12*(3), 462–496. https://doi.org/10.1044/jshr.1203.462

Darley, F. L., Aronson, A. E., & Brown, J. R. (1969b). Differential diagnostic patterns of dysarthria. *Journal of Speech and Hearing Research, 12(*2), 246–269. https://doi.org/10.1044/jshr.1202.246

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

Defebvre, L., & Vérin, M. (Éds.). (2025). *La maladie de Parkinson* (5e éd.). Elsevier Masson.

Delais-Roussarie, E., Post, B., Avanzi, M., Buthke, C., Di Cristo, A., Feldhausen, I., Jun, S.-A., Martin, P., Meisenburg, T., Rialland, A., Sichel-Bazin, R., & Yoo, H.-Y. (2015). Intonational phonology of French: Developing a ToBI system for French. In S. Frota & P. Prieto (Eds.), *Intonation in Romance* (pp. 63–100). Oxford University Press.

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review, 93*(3), 283–321. https://doi.org/10.1037/0033-295X.93.3.283

Di Cristo, A. (2004). La prosodie au carrefour de la phonétique, de la phonologie et de l'articulation formes-fonctions. *Travaux interdisciplinaires du Laboratoire Parole et Langage d'Aix-en-Provence, 23*, 67–211.

Di Cristo, A. (2013). *La prosodie de la parole*. De Boeck-Solal.

Di Cristo, A. (2016). *Les musiques du français parlé : Essais sur l'accentuation, la métrique, le rythme, le phrasé prosodique et l'intonation du français contemporain*. De Gruyter Mouton.

Duffy, J. R. (2020). *Motor speech disorders: Substrates, differential diagnosis, and management* (4th ed.). Elsevier.

Fabbri, M., Guimarães, I., Cardoso, R., Coelho, M., Guedes, L. C., Rosa, M. M., Godinho, C., Abreu, D., Gonçalves, N., Antonini, A., & Ferreira, J. J. (2017). Speech and voice response to a levodopa challenge in late-stage Parkinson's disease. *Frontiers in Neurology, 8*, 432. https://doi.org/10.3389/fneur.2017.00432

Falk, S., & Dalla Bella, S. (2016). It is better when expected: Aligning speech and motor rhythms enhances verbal processing. *Language, Cognition and Neuroscience, 31*(5), 699–708. https://doi.org/10.1080/23273798.2016.1144892

Falk, S., Volpi-Moncorger, C., & Dalla Bella, S. (2017). Auditory-motor rhythms and speech processing in French and German listeners. *Frontiers in Psychology, 8*, 395. https://doi.org/10.3389/fpsyg.2017.00395

Fearnley, J. M., & Lees, A. J. (1991). Ageing and Parkinson's disease: Substantia nigra regional selectivity. *Brain, 114*(5), 2283–2301. https://doi.org/10.1093/brain/114.5.2283

Fiveash, A., Bedoin, N., Lalitte, P., & Tillmann, B. (2020). Rhythmic priming of grammaticality judgments in children: Duration matters. *Journal of Experimental Child Psychology, 197*, 104885. https://doi.org/10.1016/j.jecp.2020.104885

Fiveash, A., Bedoin, N., Gordon, R. L., & Tillmann, B. (2021). Processing rhythm in speech and music: Shared mechanisms and implications for developmental speech and language disorders. *Neuropsychology, 35*(8), 771–791. https://doi.org/10.1037/neu0000766

Fiveash, A., Dalla Bella, S., Bigand, E., Gordon, R. L., & Tillmann, B. (2022). You got rhythm, or more: The multidimensionality of rhythmic abilities. *Attention, Perception, & Psychophysics, 84*(4), 1370–1392. https://doi.org/10.3758/s13414-022-02487-2

Fiveash, A., Bedoin, N., Gordon, R. L., & Tillmann, B. (2023). Regular rhythmic primes improve sentence repetition in children with developmental language disorder. *npj Science of Learning, 8*, 23. https://doi.org/10.1038/s41539-023-00170-1

Fiveash, A., Bedoin, N., & Tillmann, B. (2025). Examining methodological influences on the rhythmic priming effect: A commentary on Kim, McLaren, and Lee (2024). *Journal of Experimental Child Psychology, 250*, 106111. https://doi.org/10.1016/j.jecp.2024.106111

Fox, C. M., Ramig, L. O., Ciucci, M. R., Sapir, S., McFarland, D. H., & Farley, B. G. (2006). The science and practice of LSVT/LOUD: Neural plasticity-principled approach to treating

individuals with Parkinson disease and other neurological disorders. *Seminars in Speech and Language, 27*(4), 283–299. https://doi.org/10.1055/s-2006-955118

Frota, S., Cruz, M., Cardoso, R., Guimarães, I., Ferreira, J. J., Pinto, S., & Vigário, M. (2021). (Dys)Prosody in Parkinson's disease: Effects of medication and disease duration on intonation and prosodic phrasing. *Brain Sciences, 11*(8), 1100. https://doi.org/10.3390/brainsci11081100

Fumel, J., Bahuaud, D., Weed, E., Fusaroli, R., & Basirat, A. (2024). A systematic review and Bayesian meta-analysis of acoustic measures of prosody in Parkinson's disease. *Journal of Speech, Language, and Hearing Research, 67*(8), 2548–2564. https://doi.org/10.1044/2024_JSLHR-23-00588

Fujii, S., & Schlaug, G. (2013). The Harvard Beat Assessment Test (H-BAT): A battery for assessing beat perception and production and their dissociation. *Frontiers in Human Neuroscience, 7*, 771. https://doi.org/10.3389/fnhum.2013.00771

Fujii, S., & Wan, C. Y. (2014). The role of rhythm in speech and language rehabilitation: The SEP hypothesis. *Frontiers in Human Neuroscience, 8*, Article 777. https://doi.org/10.3389/fnhum.2014.00777

Galaz, Z., Mekyska, J., Mzourek, Z., Smekal, Z., Rektorova, I., Eliasova, I., Kostalova, M., Mrackova, M., & Berankova, D. (2016). Prosodic analysis of neutral, stress-modified and rhymed speech in patients with Parkinson's disease. *Computer Methods and Programs in Biomedicine, 127*, 301–317. https://doi.org/10.1016/j.cmpb.2015.12.011

Ghai, S., Ghai, I., Schmitz, G., & Effenberg, A. O. (2018). Effect of rhythmic auditory cueing on parkinsonian gait: A systematic review and meta-analysis. *Scientific Reports, 8*, 506. https://doi.org/10.1038/s41598-017-16232-5

Gindre, A.-F. (2024). *Du rythme à la parole : effet d'amorces rythmiques langagières, non langagières et musicales modulées par l'engagement moteur sur le temps d'initiation de la parole* [Thèse de doctorat, Université Toulouse–Jean Jaurès]. theses.fr. https://theses.fr/2024TLSEJ048

Goetz, C. G., Tilley, B. C., Shaftman, S. R., Stebbins, G. T., Fahn, S., Martinez-Martin, P., Poewe, W., Sampaio, C., Stern, M. B., Dodel, R., Dubois, B., Holloway, R., Jankovic, J., Kulisevsky, J., Lang, A. E., Lees, A., Leurgans, S., LeWitt, P. A., Nyenhuis, D., Olanow, C. W., Rascol, O., Schrag, A., Teresi, J. A., van Hilten, J. J., LaPelle, N., & Movement Disorder Society UPDRS Revision Task Force. (2008). Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): Scale presentation and clinimetric testing results. *Movement Disorders, 23*(15), 2129–2170. https://doi.org/10.1002/mds.22340

Goldman, J. G., Volpe, D., Ellis, T. D., Hirsch, M. A., Johnson, J., Wood, J., Aragon, A., Biundo, R., Di Rocco, A., Kasman, G. S., Iansek, R., Miyasaki, J., McConvey, V. M., Munneke, M., Pinto, S., St. Clair, K. A., Toledo, S., York, M. K., Todaro, R., Yarab, N., & Wallock, K. (2024). Delivering multidisciplinary rehabilitation care in Parkinson's disease: An international consensus statement. *Journal of Parkinson's Disease, 14*(1), 135–166. https://doi.org/10.3233/JPD-230117

Grahn, J. A., & Brett, M. (2009). Impairment of beat-based rhythm discrimination in Parkinson's disease. *Cortex, 45*(1), 54–61. https://doi.org/10.1016/j.cortex.2008.01.005

Greenwell, T., & Walsh, B. (2021). Evidence-based practice in speech-language pathology: Where are we now? *American Journal of Speech-Language Pathology, 30*(1), 186–198. https://doi.org/10.1044/2020_AJSLP-20-00194

Guenther, F. H., Ghosh, S. S., & Tourville, J. A. (2006). Neural modeling and imaging of the cortical interactions underlying syllable production. *Brain and Language, 96*(3), 280–301. https://doi.org/10.1016/j.bandl.2005.06.001

Guenther, F. H. (2016). *Neural control of speech*. MIT Press.

György, D., Saddy, J. D., Kotz, S. A., & Franck, J. (2024). Rhythmic priming of syntactic processing in Jabberwocky: A short-lived effect. *Language, Cognition and Neuroscience, 39*(7), 1–20. https://doi.org/10.1080/23273798.2024.2361737

Hartelius, L., & Svensson, P. (1994). Speech and swallowing symptoms associated with Parkinson's disease and multiple sclerosis: A survey. *Folia Phoniatrica et Logopaedica, 46*(1), 9–17. https://doi.org/10.1159/000266286

Haute Autorité de Santé. (2016). *Guide parcours de soins : Maladie de Parkinson*. https://www.has-sante.fr/jcms/c_1243427/fr/guide-parcours-de-soins-maladie-de-parkinson

Hickok, G. (2012). Computational neuroanatomy of speech production. *Nature Reviews Neuroscience, 13*(2), 135–145. https://doi.org/10.1038/nrn3158

Ho, A. K., Bradshaw, J. L., Iansek, R., & Alfredson, R. (1999). Speech volume regulation in Parkinson's disease: Effects of implicit cues and explicit instructions. *Neuropsychologia, 37(*13), 1453–1460. https://doi.org/10.1016/S0028-3932(99)00067-6

Ho, A. K., Bradshaw, J. L., & Iansek, R. (2000). Volume perception in parkinsonian speech. *Movement Disorders, 15*(6), 1125–1131. https://doi.org/10.1002/1531-8257(200011)15:6

Hornykiewicz, O. (2006). The discovery of dopamine deficiency in the parkinsonian brain. *Journal of Neural Transmission. Supplementum, 70*, 9–15. https://doi.org/10.1007/978-3-211-45295-0_3

Houde, J. F., & Nagarajan, S. S. (2011). Speech production as state feedback control. *Frontiers in Human Neuroscience, 5,* 82. https://doi.org/10.3389/fnhum.2011.00082

Jones, M. R., & Boltz, M. (1989). Dynamic attending and responses to time. *Psychological Review, 96*(3), 459–491. https://doi.org/10.1037/0033-295X.96.3.459

Jones, M. R. (2019). *Time will tell: A theory of dynamic attending.* Oxford University Press. https://doi.org/10.1093/oso/9780190618216.001.0001

Jun, S.-A., & Fougeron, C. (2000). A phonological model of French intonation. In A. Botinis (Ed.), *Intonation: Analysis, modelling and technology* (pp. 209–242). Kluwer Academic Publishers. https://doi.org/10.1007/978-94-011-4317-2_10

Jun, S.-A., & Fougeron, C. (2002). Realizations of accentual phrase in French intonation. *Probus, 14*(1), 147–172. https://doi.org/10.1515/prbs.2002.002

Jungers, M. K., & Hupp, J. M. (2018). Music to my mouth: Evidence of domain general rate priming in adults and children. *Cognitive Development, 48*, 219–224. https://doi.org/10.1016/j.cogdev.2018.09.001

Kent, R. D., Kent, J. F., Duffy, J. R., Thomas, J. E., Weismer, G., & Stuntebeck, S. (2000). Ataxic dysarthria. *Journal of Speech, Language, and Hearing Research, 43*(5), 1275–1289. https://doi.org/10.1044/jslhr.4305.1275

Kleiner, M., Brainard, D., & Pelli, D. (2007). What's new in Psychtoolbox-3? *Perception, 36(ECVP Abstract Supplement), 14*. https://doi.org/10.1177/03010066070360S101

Kösem, A., Bosker, H. R., Takashima, A., Meyer, A., Jensen, O., & Hagoort, P. (2018). Neural entrainment determines the words we hear. *Current Biology, 28*(18), 2867–2875.e3. https://doi.org/10.1016/j.cub.2018.07.023

Koshimori, Y., & Thaut, M. H. (2018). Future perspectives on neural mechanisms underlying rhythm and music based neurorehabilitation in Parkinson's disease. *Ageing Research Reviews, 47*, 133–139. https://doi.org/10.1016/j.arr.2018.07.001

Koshimori, Y., & Thaut, M. H. (2023). Rhythmic auditory stimulation as a potential neuromodulator for Parkinson's disease. *Parkinsonism & Related Disorders, 113*, 105459. https://doi.org/10.1016/j.parkreldis.2023.105459

Kotz, S. A., & Gunter, T. C. (2015). Can rhythmic auditory cuing remediate language-related deficits in Parkinson's disease? *Annals of the New York Academy of Sciences, 1337*(1), 62–68. https://doi.org/10.1111/nyas.12657

Kotz, S. A., & Schwartze, M. (2016). Motor-timing and sequencing in speech production: A general-purpose framework. In G. Hickok & S. L. Small (Eds.), *Neurobiology of language* (pp. 717–724). Academic Press. https://doi.org/10.1016/B978-0-12-407794-2.00057-2

Kraus, N., & Chandrasekaran, B. (2010). Music training for the development of auditory skills. *Nature Reviews Neuroscience, 11*(8), 599–605. https://doi.org/10.1038/nrn2882

Kurz, A. S. (2019, February 10). *Bayesian robust correlations with brms (and why you should love Student's t)*. A. Solomon Kurz. https://solomonkurz.netlify.app/blog/2019-02-10-bayesian-robust-correlations-with-brms-and-why-you-should-love-student-s-t/

Lacheret, A. (2011). La prosodie au cœur du verbal. *Rééducation Orthophonique, 246,* 87–104.

Ladányi, E., Lukács, Á., & Gervain, J. (2021). Does rhythmic priming improve grammatical processing in Hungarian-speaking children with and without developmental language disorder? *Developmental Science, 24*, e13112. https://doi.org/10.1111/desc.13112

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

Large, E. W., & Jones, M. R. (1999). The dynamics of attending: How people track time-varying events. *Psychological Review, 106*(1), 119–159. https://doi.org/10.1037/0033-295X.106.1.119

Law, L. N. C., & Zentner, M. (2012). Assessing musical abilities objectively: Construction and validation of the Profile of Music Perception Skills. *PLOS ONE, 7*(12), e52508. https://doi.org/10.1371/journal.pone.0052508

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences, 22*(1), 1–38. https://doi.org/10.1017/S0140525X99001776

Levy, E. S., Moya-Galé, G., Chang, Y. H. M., Freeman, K., Forrest, K., Brin, M. F., & Ramig, L. A. (2020). The effects of intensive speech treatment on intelligibility in Parkinson's disease: A randomised controlled trial. *EClinicalMedicine, 24*, 100429. https://doi.org/10.1016/j.eclinm.2020.100429

L'Hermite, S., & Zoefel, B. (2023). Rhythmic entrainment echoes in auditory perception. *Journal of Neuroscience, 43*(39), 6667–6678. https://doi.org/10.1523/JNEUROSCI.0051-23.2023

Lim, I., van Wegen, E., de Goede, C., Deutekom, M., Nieuwboer, A., Willems, A.-M., Jones, D., Rochester, L., & Kwakkel, G. (2005). Effects of external rhythmical cueing on gait in patients with Parkinson's disease: A systematic review. *Clinical Rehabilitation, 19*(7), 695–713. https://doi.org/10.1191/0269215505cr906oa

Liss, J. M., White, L., Mattys, S. L., Lansford, K., Lotto, A. J., Spitzer, S. M., & Caviness, J. N. (2009). Quantifying speech rhythm abnormalities in the dysarthrias. *Journal of Speech, Language, and Hearing Research, 52*(5), 1334–1352. https://doi.org/10.1044/1092-4388(2009/08-0208)

Liu, X., & Liu, Y. (2024). Music rhythmic cueing for the production of non-native speech rhythm: Evidence from Chinese learners of French. *Journal of Psycholinguistic Research, 53*, 10. https://doi.org/10.1007/s10936-024-10044-1

Lizcano-Cortés, F., Gómez-Varela, I., Mares, C., Wallisch, P., Orpella, J., Poeppel, D., Ripollés, P., & Assaneo, M. F. (2022). Speech-to-speech synchronization protocol to classify human participants as high or low auditory-motor synchronizers. *STAR Protocols, 3*(2), Article 101248. https://doi.org/10.1016/j.xpro.2022.101248

Lloyd, A. J. (1999). Comprehension of prosody in Parkinson's disease. *Cortex, 35*(3), 389–402. https://doi.org/10.1016/S0010-9452(08)70807-4

Lowit, A. (2014). Quantification of rhythm problems in disordered speech: A re-evaluation. *Philosophical Transactions of the Royal Society B: Biological Sciences, 369*(1658), 20130404. https://doi.org/10.1098/rstb.2013.0404

Lowit, A., & Kuschmann, A. (2012). Characterizing intonation deficit in motor speech disorders: An autosegmental–metrical analysis of spontaneous speech in hypokinetic dysarthria, ataxic dysarthria, and foreign accent syndrome. *Journal of Speech, Language, and Hearing Research, 55*(5), 1472–1484. https://doi.org/10.1044/1092-4388(2012/11-0263)

Manes, J. L., Bullock, L., Meier, A. M., Turner, R. S., Richardson, R. M., & Guenther, F. H. (2024). A neurocomputational view of the effects of Parkinson's disease on speech production. *Frontiers in Human Neuroscience, 18*, Article 1383714. https://doi.org/10.3389/fnhum.2024.1383714

Manes, J. L., Stipancic, K. L., Westemeyer, L., Rohl, A., Greenlee, J. D. W., Narayanan, N., Uc, E., & Tjaden, K. (2026). Perceptual outcome measures for supplementing transcription intelligibility in speakers with Parkinson's disease and mild dysarthria. *Folia Phoniatrica et Logopaedica. Advance online publication.* https://doi.org/10.1159/000550174

Mariën, P., Ackermann, H., Adamaszek, M., Barwood, C. H. S., Beaton, A., Desmond, J., De Witte, E., Fawcett, A. J., Hertrich, I., Küper, M., Leggio, M., Marvel, C., Molinari, M., Murdoch, B. E., Nicolson, R. I., Schmahmann, J. D., Stoodley, C. J., Thürling, M., Timmann, D., Wouters, E., & Ziegler, W. (2014). Consensus paper: Language and the cerebellum: An ongoing enigma. *The Cerebellum, 13*(3), 386–410. https://doi.org/10.1007/s12311-013-0540-5

Martens, H., Van Nuffelen, G., Dekens, T., Hernández-Díaz Huici, M. E., Kairuz Hernández-Díaz, H. A., De Letter, M., & De Bodt, M. (2015). The effect of intensive speech rate and intonation therapy on intelligibility in Parkinson's disease. *Journal of Communication Disorders, 58*, 91–105. https://doi.org/10.1016/j.jcomdis.2015.10.004

McGregor, M. M., & Nelson, A. B. (2019). Circuit mechanisms of Parkinson's disease. *Neuron, 101*(6), 1042–1056. https://doi.org/10.1016/j.neuron.2019.03.004

McIntosh, G. C., Brown, S. H., Rice, R. R., & Thaut, M. H. (1997). Rhythmic auditory-motor facilitation of gait patterns in patients with Parkinson's disease. *Journal of Neurology, Neurosurgery & Psychiatry, 62*(1), 22–26. https://doi.org/10.1136/jnnp.62.1.22

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

Miller, N., Noble, E., Jones, D., & Burn, D. (2006). Life with communication changes in Parkinson's disease. *Age and Ageing, 35*(3), 235–239. https://doi.org/10.1093/ageing/afj053

Miller, N. (2017). Communication changes in Parkinson's disease. *Practical Neurology, 17*(4), 266–274. https://doi.org/10.1136/practneurol-2017-001635

Mollaei, F., Evans, H., & Pool, A. (2026). Pause characteristics of sentence production in Parkinson's disease: Insights from sentence complexity and length. *PLOS ONE, 21*(4), Article e0330808. https://doi.org/10.1371/journal.pone.0330808

Moreau, C., Pennel-Ployart, O., Pinto, S., Plachez, A., Annic, A., Viallet, F., Destée, A., & Defebvre, L. (2011). Modulation of dysarthropneumophonia by low-frequency STN DBS in advanced Parkinson's disease. *Movement Disorders, 26*(4), 659–663. https://doi.org/10.1002/mds.23538

Moreau, C., & Pinto, S. (2019). Misconceptions about speech impairment in Parkinson's disease. *Movement Disorders, 34*(10), 1471–1475. https://doi.org/10.1002/mds.27791

Moya-Galé, G., Levy, E. S., Chang, Y. H. M., & Ramig, L. A. (2024). ¿Notas la diferencia? [Do you hear the difference?]: Perceptual consequences of intensive voice treatment in Spanish speakers with Parkinson's disease. *Journal of Speech, Language, and Hearing Research, 67*(9), 2872–2892. https://doi.org/10.1044/2023_JSLHR-23-00379

Muñoz-Vigueras, N., Prados-Román, E., Valenza, M. C., Granados-Santiago, M., Cabrera-Martos, I., Rodríguez-Torres, J., & Torres-Sánchez, I. (2021). Speech and language therapy treatment on hypokinetic dysarthria in Parkinson disease: Systematic review and meta-analysis. *Clinical Rehabilitation, 35*(5), 639–655. https://doi.org/10.1177/0269215520976267

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

Nieuwboer, A., Kwakkel, G., Rochester, L., Jones, D., van Wegen, E., Willems, A.-M., Chavret, F., Hetherington, V., Baker, K., & Lim, I. (2007). Cueing training in the home improves gait-related mobility in Parkinson's disease: The RESCUE trial. *Journal of Neurology, Neurosurgery & Psychiatry, 78*(2), 134–140. https://doi.org/10.1136/jnnp.200X.097923

Nombela, C., Hughes, L. E., Owen, A. M., & Grahn, J. A. (2013). Into the groove: Can rhythm influence Parkinson's disease? *Neuroscience & Biobehavioral Reviews, 37*(10), 2564–2570. https://doi.org/10.1016/j.neubiorev.2013.08.003

Nozaradan, S., Schwartze, M., Obermeier, C., & Kotz, S. A. (2017). Specific contributions of basal ganglia and cerebellum to the neural tracking of rhythm. *Cortex, 95*, 156–168. https://doi.org/10.1016/j.cortex.2017.08.015

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia, 9*(1), 97–113. https://doi.org/10.1016/0028-3932(71)90067-4

Ong, A., Namasivayam-MacDonald, A., Kim, S., & Werden Abrams, S. (2024). The use of music and music-related elements in speech-language therapy interventions for adults with neurogenic communication impairments: A scoping review. *International Journal of Language & Communication Disorders, 59*(6), 2632–2654. https://doi.org/10.1111/1460-6984.13104

Parrell, B., Ramanarayanan, V., Nagarajan, S., & Houde, J. (2019). The FACTS model of speech motor control: Fusing state estimation and task-based control. *PLOS Computational Biology, 15*(9), e1007321. https://doi.org/10.1371/journal.pcbi.1007321

Patel, A. D. (2011). Why would musical training benefit the neural encoding of speech? The OPERA hypothesis. *Frontiers in Psychology, 2*, 142. https://doi.org/10.3389/fpsyg.2011.00142

Patel, A. D. (2012). The OPERA hypothesis: Assumptions and clarifications. *Annals of the New York Academy of Sciences, 1252*(1), 124–128. https://doi.org/10.1111/j.1749-6632.2011.06426.x

Patel, A. D., & Iversen, J. R. (2014). The evolutionary neuroscience of musical beat perception: The Action Simulation for Auditory Prediction (ASAP) hypothesis. *Frontiers in Systems Neuroscience, 8*, 57. https://doi.org/10.3389/fnsys.2014.00057

Peelle, J. E., & Davis, M. H. (2012). Neural oscillations carry speech rhythm through to comprehension. *Frontiers in Psychology, 3*, 320. https://doi.org/10.3389/fpsyg.2012.00320

Pell, M. D. (1996). On the receptive prosodic loss in Parkinson's disease. *Cortex, 32*(4), 693–704. https://doi.org/10.1016/S0010-9452(96)80039-6

Pell, M. D., Cheang, H. S., & Leonard, C. L. (2006). The impact of Parkinson's disease on vocal-prosodic communication from the perspective of listeners. *Brain and Language, 97*(2), 123–134. https://doi.org/10.1016/j.bandl.2005.08.010

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10*(4), 437–442. https://doi.org/10.1163/156856897X00366

Pernon, M., Levêque, N., Delvaux, V., Assal, F., Borel, S., Fougeron, C., Trouville, R., & Laganaro, M. (2020). MonPaGe, un outil de screening francophone informatisé d'évaluation des troubles moteurs de la parole. *Rééducation Orthophonique, 281*, 171–197.

Petitpain, A., Brabant-Thery, J., Patin, C., Ott, L., & Basirat, A. (2024). Status of prosody in the practices of speech-language pathologists in France: A survey on theoretical and clinical dimensions. *Journal of Communication Disorders, 107*, 106404. https://doi.org/10.1016/j.jcomdis.2023.106404

Pinto, S., Gentil, M., Krack, P., Sauleau, P., Fraix, V., Benabid, A.-L., & Pollak, P. (2005). Changes induced by levodopa and subthalamic nucleus stimulation on parkinsonian speech. Movement Disorders, 20(11), 1507–1515. https://doi.org/10.1002/mds.20601

Pinto, S., Ghio, A., Teston, B., & Viallet, F. (2010). La dysarthrie au cours de la maladie de Parkinson. Histoire naturelle de ses composantes : dysphonie, dysprosodie et dysarthrie. *Revue Neurologique, 166*(10), 800–810. https://doi.org/10.1016/j.neurol.2010.07.005

Pinto, S., Cardoso, R., Sadat, J., Guimarães, I., Mercier, C., Santos, H., Atkinson-Clement, C., Carvalho, J., Welby, P., Oliveira, P., D'Imperio, M., Frota, S., Letanneux, A., Vigário, M., Cruz, M., Martins, I. P., Viallet, F., & Ferreira, J. J. (2016). Dysarthria in individuals with Parkinson's disease: A protocol for a binational, cross-sectional, case-controlled study in French and

European Portuguese (FraLusoPark). *BMJ Open, 6*(11), e012885. https://doi.org/10.1136/bmjopen-2016-012885

Pinto, S., Nebel, A., Rau, J., Espesser, R., Maillochon, P., Niebuhr, O., Krack, P., Witjas, T., Ghio, A., Cuartero, M.-C., Timmermann, L., Schnitzler, A., Hesekamp, H., Meier, N., Müllner, J., Hälbig, T. D., Möller, B., Paschen, S., Paschen, L., ... Deuschl, G. (2023). Results of a randomized clinical trial of speech after early neurostimulation in Parkinson's disease. *Movement Disorders, 38*(2), 212–222. https://doi.org/10.1002/mds.29282

Prenger, M. T. M., Madray, R., Van Hedger, K., Anello, M., & MacDonald, P. A. (2020). Social symptoms of Parkinson's disease. *Parkinson's Disease, 2020*, 8846544. https://doi.org/10.1155/2020/8846544

Przybylski, L., Bedoin, N., Krifi-Papoz, S., Herbillon, V., Roch, D., Léculier, L., Kotz, S. A., & Tillmann, B. (2013). Rhythmic auditory stimulation influences syntactic processing in children with developmental language disorders. *Neuropsychology, 27*(1), 121–131. https://doi.org/10.1037/a0031277

Pu, T., Huang, M., Kong, X., Wang, M., Chen, X., Feng, X., Wei, C., & Yang, J. (2021). Lee Silverman Voice Treatment to improve speech in Parkinson's disease: A systemic review and meta-analysis. *Parkinson's Disease, 2021*, 3366870. https://doi.org/10.1155/2021/3366870

Puyjarinet, F., Bégel, V., Gény, C., Driss, V., Cuartero, M.-C., Kotz, S. A., Pinto, S., & Dalla Bella, S. (2019). Heightened orofacial, manual, and gait variability in Parkinson's disease results from a general rhythmic impairment. *npj Parkinson's Disease, 5*, 19. https://doi.org/10.1038/s41531-019-0092-6

Puyjarinet, F., Bégel, V., Gény, C., Driss, V., Cuartero, M.-C., Cochen De Cock, V., Pinto, S., & Dalla Bella, S. (2022). At-home training with a rhythmic video game for improving orofacial, manual, and gait abilities in Parkinson's disease: A pilot study. *Frontiers in Neuroscience, 16*, 874032. https://doi.org/10.3389/fnins.2022.874032

Ramig, L. O., Sapir, S., Countryman, S., Pawlas, A. A., O'Brien, C., Hoehn, M., & Thompson, L. L. (2001). Intensive voice treatment (LSVT) for patients with Parkinson's disease: A 2 year follow up. *Journal of Neurology, Neurosurgery & Psychiatry, 71*(4), 493–498. https://doi.org/10.1136/jnnp.71.4.493

Repp, B. H. (2005). Sensorimotor synchronization: A review of the tapping literature. *Psychonomic Bulletin & Review, 12*(6), 969–992. https://doi.org/10.3758/BF03206433

Rusz, J., Čmejla, R., Růžičková, H., & Růžička, E. (2011). Quantitative acoustic measurements for characterization of speech and voice disorders in early untreated Parkinson's disease. *The Journal of the Acoustical Society of America, 129*(1), 350–367. https://doi.org/10.1121/1.3514381

Rusz, J., Hlavnička, J., Novotný, M., Tykalová, T., Pelletier, A., Montplaisir, J., Gagnon, J.-F., Dušek, P., Galbiati, A., Marelli, S., Timm, P. C., Teigen, L. N., Janzen, A., Habibi, M., Stefani, A., Holzknecht, E., Seppi, K., Evangelista, E., Rassu, A. L., Dauvilliers, Y., Högl, B., Oertel, W., St. Louis, E. K., Ferini-Strambi, L., Růžička, E., Postuma, R. B., & Šonka, K. (2021a). Speech

biomarkers in rapid eye movement sleep behavior disorder and Parkinson disease. *Annals of Neurology, 90*(1), 62–75. https://doi.org/10.1002/ana.26085

Rusz, J., Tykalová, T., Novotný, M., Zogala, D., Šonka, K., Růžička, E., & Dušek, P. (2021b). Defining speech subtypes in de novo Parkinson disease: Response to long-term levodopa therapy. *Neurology, 97*(21), e2124–e2135. https://doi.org/10.1212/WNL.0000000000012878

Rusz, J., Krack, P., & Tripoliti, E. (2024). From prodromal stages to clinical trials: The promise of digital speech biomarkers in Parkinson's disease. *Neuroscience & Biobehavioral Reviews, 167*, 105922. https://doi.org/10.1016/j.neubiorev.2024.105922

Sackley, C. M., Rick, C., Au, P., Brady, M. C., Carragher, M., Hewer, R., Jowett, S., Patel, S., Smith, C. H., Deane, K. H. O., Woolley, R., Clarke, C. E., & PD COMM Collaborative Group. (2024). Lee Silverman voice treatment versus NHS speech and language therapy versus control for dysarthria in people with Parkinson's disease (PD COMM): Pragmatic, UK based, multicentre, three arm, parallel group, unblinded, randomised controlled trial. *BMJ, 386*, e078341. https://doi.org/10.1136/bmj-2023-078341

Saliba, J., Lorenzo-Seva, U., Marco-Pallarés, J., Tillmann, B., Zeitouni, A., & Lehmann, A. (2016). French validation of the Barcelona Music Reward Questionnaire. *PeerJ, 4*, e1760. https://doi.org/10.7717/peerj.1760

Santé publique France. (2023, 12 avril). Données : Maladie de Parkinson. https://www.santepubliquefrance.fr/maladie-de-parkinson/donnees

Sapir, S., Spielman, J. L., Ramig, L. O., Story, B. H., & Fox, C. (2007). Effects of intensive voice treatment (the Lee Silverman Voice Treatment [LSVT]) on vowel articulation in dysarthric individuals with idiopathic Parkinson disease: Acoustic and perceptual findings. *Journal of Speech, Language, and Hearing Research, 50*(4), 899–912. https://doi.org/10.1044/1092-4388(2007/064)

Schalling, E., & Hartelius, L. (2013). Speech in spinocerebellar ataxia. *Brain and Language, 127*(3), 317–322. https://doi.org/10.1016/j.bandl.2013.10.002

Schalling, E., Johansson, K., & Hartelius, L. (2017). Speech and communication changes reported by people with Parkinson's disease. *Folia Phoniatrica et Logopaedica, 69*(3), 131–141. https://doi.org/10.1159/000479927

Schulz, G. M., & Grant, M. K. (2000). Effects of speech therapy and pharmacologic and surgical treatments on voice and speech in Parkinson's disease: A review of the literature. *Journal of Communication Disorders, 33*(1), 59–88. https://doi.org/10.1016/S0021-9924(99)00025-8

Schwartz, R., & Pell, M. D. (2012). Emotional speech processing at the intersection of prosody and semantics. *PLOS ONE, 7*(10), e47279. https://doi.org/10.1371/journal.pone.0047279

Simon, A. C., Avanzi, M., & Goldman, J.-P. (2008). La détection des proéminences syllabiques : Un aller-retour entre l'annotation manuelle et le traitement automatique. In *Actes du CMLF 2008 – 1er Congrès Mondial de Linguistique Française* (pp. 1685–1698). https://doi.org/10.1051/cmlf08256

Skodda, S. (2012). Effect of deep brain stimulation on speech performance in Parkinson's disease. *Parkinson's Disease, 2012*, 850596. https://doi.org/10.1155/2012/850596

Späth, M., Aichert, I., Ceballos-Baumann, A. O., Wagner-Sonntag, E., Miller, N., & Ziegler, W. (2016). Entraining with another person's speech rhythm: Evidence from healthy speakers and individuals with Parkinson's disease. *Clinical Linguistics & Phonetics, 30*(1), 68–85. https://doi.org/10.3109/02699206.2015.1115129

Späth, M., Aichert, I., Timmann, D., Ceballos-Baumann, A. O., Wagner-Sonntag, E., & Ziegler, W. (2022). The role of the basal ganglia and cerebellum in adaptation to others' speech rate and rhythm: A study of patients with Parkinson's disease and cerebellar degeneration. *Cortex, 157*, 81–98. https://doi.org/10.1016/j.cortex.2022.08.012

de Souza, G. B., & Beber, B. C. (2026). Language skills in Parkinson's disease: State of the art. *Dementia & Neuropsychologia, 20,* e20250350. https://doi.org/10.1590/1980-5764-DN-2025-0350

Stockbridge, M. D., Sheppard, S. M., Keator, L. M., Murray, L. L., Lehman Blake, M., & Right Hemisphere Disorders Working Group, Evidence-Based Clinical Research Committee, Academy of Neurological Communication Disorders and Sciences. (2022). Aprosodia subsequent to right hemisphere brain damage: A systematic review and meta-analysis. *Journal of the International Neuropsychological Society, 28*(7), 709–735. https://doi.org/10.1017/S1355617721000825

Strauss, H., Reiche, S., Dick, M., & Zentner, M. (2024). Online assessment of musical ability in 10 minutes: Development and validation of the Micro-PROMS. *Behavior Research Methods, 56*(3), 1968–1983. https://doi.org/10.3758/s13428-023-02130-4

Teki, S., Grube, M., Kumar, S., & Griffiths, T. D. (2011). Distinct neural substrates of duration-based and beat-based auditory timing. *The Journal of Neuroscience, 31*(10), 3805–3812. https://doi.org/10.1523/JNEUROSCI.5561-10.2011

Teki, S., Grube, M., & Griffiths, T. D. (2012). A unified model of time perception accounts for duration-based and beat-based timing mechanisms. *Frontiers in Integrative Neuroscience, 5*, 90. https://doi.org/10.3389/fnint.2011.00090

Thaut, M. H., McIntosh, G. C., Rice, R. R., Miller, R. A., Rathbun, J., & Brault, J. M. (1996). Rhythmic auditory stimulation in gait training for Parkinson's disease patients. *Movement Disorders, 11*(2), 193–200. https://doi.org/10.1002/mds.870110213

Thaut, M. H., McIntosh, K. W., McIntosh, G. C., & Hoemberg, V. (2001). Auditory rhythmicity enhances movement and speech motor control in patients with Parkinson's disease. *Functional Neurology, 16*(2), 163–172.

Thaut, M. H., & Abiru, M. (2010). Rhythmic auditory stimulation in rehabilitation of movement disorders: A review of current research. *Music Perception, 27*(4), 263–269. https://doi.org/10.1525/mp.2010.27.4.263

Thies, T., Mücke, D., Lowit, A., Kalbe, E., Steffen, J., & Barbe, M. T. (2020). Prominence marking in Parkinsonian speech and its correlation with motor performance and cognitive abilities. *Neuropsychologia, 137*, 107306. https://doi.org/10.1016/j.neuropsychologia.2019.107306

Tierney, A., & Kraus, N. (2013). The ability to tap to a beat relates to cognitive, linguistic, and perceptual skills. *Brain and Language, 124*(3), 225–231. https://doi.org/10.1016/j.bandl.2012.12.014

Tinaz, S., Pillai, A. S., & Hallett, M. (2016). Sequence effect in Parkinson's disease is related to motor energetic cost. Frontiers in *Neurology, 7*, 83. https://doi.org/10.3389/fneur.2016.00083

Tjaden, K., & Wilding, G. (2011). Effects of speaking task on intelligibility in Parkinson's disease. Clinical *Linguistics & Phonetics, 25*(2), 155–168. https://doi.org/10.3109/02699206.2010.520185

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

Tripoliti, E., Zrinzo, L., Martinez-Torres, I., Frost, E., Pinto, S., Foltynie, T., Holl, E., Petersen, E., Roughton, M., Hariz, M. I., & Limousin, P. (2011). Effects of subthalamic stimulation on speech of consecutive patients with Parkinson disease. *Neurology, 76*(1), 80–86. https://doi.org/10.1212/WNL.0b013e318203e7d0

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

Tseng, Y., Diedrichsen, J., Krakauer, J. W., Shadmehr, R., & Bastian, A. J. (2007). Sensory prediction errors drive cerebellum-dependent adaptation of reaching. *Journal of Neurophysiology, 98*(1), 54–62. https://doi.org/10.1152/jn.00266.2007

Turk, A., & Shattuck-Hufnagel, S. (2020). *Speech timing: Implications for theories of phonology, phonetics, and speech motor control*. Oxford University Press.

Véron-Delor, L. (2021). *Évaluation et rééducation des troubles de l'écriture : vers la sonification musicale de la dysgraphie dans la maladie de Parkinson* [Thèse de doctorat, Aix-Marseille Université]. theses.fr. https://theses.fr/2021AIXM0468

Véron-Delor, L., Pinto, S., Eusebio, A., Azulay, J.-P., Witjas, T., Velay, J.-L., & Danna, J. (2020). Musical sonification improves motor control in Parkinson's disease: A proof of concept with handwriting. *Annals of the New York Academy of Sciences, 1465*(1), 132–145. https://doi.org/10.1111/nyas.14252

Weed, E., & Fusaroli, R. (2020). Acoustic measures of prosody in right-hemisphere damage: A systematic review and meta-analysis. *Journal of Speech, Language, and Hearing Research, 63*(6), 1762–1775. https://doi.org/10.1044/2020_JSLHR-19-00241

Ye, X., Li, L., He, R., Jia, Y., & Poon, W. (2022). Rhythmic auditory stimulation promotes gait recovery in Parkinson's patients: A systematic review and meta-analysis. *Frontiers in Neurology, 13*, 940419. https://doi.org/10.3389/fneur.2022.940419

Yorkston, K., Baylor, C., & Britton, D. (2017). Speech versus speaking: The experiences of people with Parkinson's disease and implications for intervention. *American Journal of Speech-Language Pathology, 26*(2S), 561–568. https://doi.org/10.1044/2017_AJSLP-16-0087

Zhang, N., & Zhang, Q. (2019). Rhythmic pattern facilitates speech production: An ERP study. *Scientific Reports, 9*, 12974. https://doi.org/10.1038/s41598-019-49375-8

Zhu, J., Cui, Y., Zhang, J., Yan, R., Su, D., Zhao, D., Wang, A., & Feng, T. (2024). Temporal trends in the prevalence of Parkinson's disease from 1980 to 2023: A systematic review and meta-analysis. *The Lancet Healthy Longevity, 5*(7), e464–e479. https://doi.org/10.1016/S2666-7568(24)00094-1

Ziegler, W., & Wessel, K. (1996). Speech timing in ataxic disorders: Sentence production and rapid repetitive articulation. *Neurology, 47*(1), 208–214. https://doi.org/10.1212/WNL.47.1.208

Zoefel, B. (2018). Speech entrainment: Rhythmic predictions carried by neural oscillations. *Current Biology, 28*(18), R1102–R1104. https://doi.org/10.1016/j.cub.2018.07.048

Zoefel, B., Abbasi, O., Gross, J., & Kotz, S. A. (2024). Entrainment echoes in the cerebellum. *Proceedings of the National Academy of Sciences of the United States of America, 121*(34), e2411167121. https://doi.org/10.1073/pnas.2411167121