

Doctorat de l'Université de Toulouse

préparé à l'Université Toulouse - Jean Jaurès

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 présentée et soutenue, le 26 septembre 2024 par

Anne-Flore GINDRE

École doctorale

CLESCO - Comportement, Langage, Éducation, Socialisation, Cognition

Spécialité

Sciences du langage

Unité de recherche

LNPL - Laboratoire de NeuroPsychoLinguistique

Thèse dirigée par

Corine ASTESANO et Halima SAHRAOUI

Composition du jury

Mme Hélène LÖVENBRUCK, Présidente, Université Grenoble Alpes

Mme Véronique BOULENGER, Rapporteur, CNRS Rhône Auvergne

Mme Christelle DODANE, Rapporteur, Université Sorbonne Nouvelle

M. Jérémy DANNA, Examineur, Université Toulouse - Jean Jaurès

Mme Corine ASTESANO, Directrice de thèse, Université Toulouse - Jean Jaurès

Mme Halima SAHRAOUI, Co-directrice de thèse, Université Toulouse - Jean Jaurès

Remerciements

Mes premiers remerciements s'adressent aux membres de mon jury de soutenance, Véronique Boulenger, Christelle Dodane, Jérémy Danna et Hélène Løevenbruck, qui ont accepté d'évaluer et examiner mon travail de thèse. Je les remercie pour leur considération et leur souhaite une belle lecture !

En second lieu, je remercie mes directrices de thèse, Corine Astésano et Halima Sahraoui, qui m'ont soutenue avant même que le projet de thèse n'existe et ont couru jusqu'aux derniers mètres de ce marathon à mes côtés. Je les remercie pour leur temps, leur investissement, leurs précieux conseils.

Je remercie également les membres de mon comité de suivi de thèse, Régis Missire et Anne-Catherine Simon, pour leurs suggestions pertinentes et nos discussions.

Je remercie Julien Tardieu, de la plateforme CCU, qui m'a fait confiance en me laissant utiliser le plateau technique PETRA comme salle d'expérience et m'a accompagnée dans la finalisation de mon protocole.

Je remercie les pré-testeurs et participants de mon étude, pour leur temps et leur intérêt motivant.

Je remercie également le Laboratoire de NeuroPsychoLinguistique qui m'a accueillie durant ces cinq années, et tout particulièrement Cyrille Granget et Stéphanie Saleilles, qui ont toujours répondu à tout et une attention envers chacune et chacun.

Mes remerciements s'adressent aussi à celles et ceux qui m'ont épaulée tout au long de cette aventure, notamment par leur soutien méthodologique (Lucille Soulier, Nour Ezzeddine, Cyril Perret, Karla Orihuela, Dimitri Paisios, Valentin Goulette, Marie Hevrova) et statistique (Clara Solier, Fanny Guitard-Ivent, Gwenaël Kaminski, Quentin Marre, Clément Steiner, Anaïs Desbernats, Salomé Gindre, Noémie Te Rietmolen).

Je remercie également la team Voltaire+ pour ses relectures attentives, conseils et encouragements, Nour, Mélanie, Marion, Alexandre, Boulou, Lara, Yoyo et mes parents, Cyrille et Marie.

Je remercie enfin la joyeuse équipe de doctorants, actuels ou anciens, toujours MDR, qui a partagé mon quotidien, mes pauses ou mes matchs de badminton, mais surtout pour tout son soutien, Mélanie, Nour, Marie, Romain, Lyanne, Chloé, Lucie, Jasmin, Lucille, Sangho, Mireille, Noémie, Katia, Layana, Kleo, Cécile, Stéphanie, Victoria, Iona et les nouvelles.

TABLE DES MATIERES

| | |
|---|-----|
| INTRODUCTION | 1 |
| I. PARTIE THEORIQUE | 5 |
| 1 UNE PAROLE RYTHMIQUE | 6 |
| 1.1 QU'EST-CE QUE LE RYTHME ? | 6 |
| 1.2 LE RYTHME DANS LA PAROLE | 13 |
| 1.3 SYNTHESE : UNE PAROLE RYTHMIQUE | 35 |
| 2 LE ROLE DU RYTHME DANS LE TRAITEMENT DE LA PAROLE | 37 |
| 2.1 MECANISMES DE TRAITEMENT DE LA PAROLE | 37 |
| 2.2 ÊTRE SUR LA MEME LONGUEUR D'ONDE, UN PREREQUIS POUR LA COMMUNICATION | 49 |
| 2.3 LANGAGE ET SUPRAMODALITE DU RYTHME | 61 |
| 2.4 SYNTHESE : DES STRUCTURES RYTHMIQUES A LEUR TRAITEMENT NEUROPSYCHOLINGUISTIQUE EN PAROLE | 70 |
| II. PROBLEMATIQUE ET HYPOTHESES | 74 |
| III. METHODOLOGIE EXPERIMENTALE | 84 |
| 3 PLAN EXPERIMENTAL | 86 |
| 3.1 PLAN EXPERIMENTAL DE L'ACTIVITE 1 : CARACTERISATION DES CAPACITES DE TRAITEMENT DU RYTHME | 86 |
| 3.2 PLAN EXPERIMENTAL DE L'ACTIVITE 2 : AMORÇAGE RYTHMIQUE DE LA PAROLE | 89 |
| 4 PROCEDURE | 104 |
| 4.1 CONSIDERATIONS ETHIQUES ET ANONYMISATION DES DONNEES | 104 |
| 4.2 MATERIEL TECHNIQUE | 104 |
| 4.3 DEROULE | 106 |
| 5 POPULATION | 109 |
| 5.1 CRITERES DE SELECTION DES PARTICIPANTS | 109 |
| 5.2 RECRUTEMENT | 110 |
| 5.3 ESTIMATION DU NOMBRE DE PARTICIPANTS | 111 |
| 5.4 DESCRIPTION DE L'ECHANTILLON | 111 |

| | | |
|------------------------------------|--|-----|
| 6 | DONNEES RECUEILLIES | 113 |
| 6.1 | DONNEES DU QUESTIONNAIRE SOCIO-DEMOGRAPHIQUE..... | 113 |
| 6.2 | DONNEES DE L'ACTIVITE 1 : CARACTERISATION DES CAPACITES DE TRAITEMENT DU RYTHME..... | 114 |
| 6.3 | DONNEES DE L'ACTIVITE 2 : AMORÇAGE RYTHMIQUE DE LA PAROLE..... | 119 |
| 7 | DEMARCHE STATISTIQUE..... | 124 |
| 7.1 | MODELE MIXTE | 124 |
| IV. RESULTATS | | 130 |
| 8 | ACTIVITE 1 : CARACTERISATION DES CAPACITES DE TRAITEMENT DU RYTHME | 133 |
| 8.1 | TACHE DE TEMPO MOTEUR SPONTANE | 134 |
| 8.2 | TACHE DE SYNCHRONISATION SENSORI-MOTRICE..... | 145 |
| 9 | ACTIVITE 2 : AMORÇAGE RYTHMIQUE DE LA PAROLE | 169 |
| 9.1 | ERREURS DE LECTURE | 171 |
| 9.2 | TEMPS D'INITIATION DE LA PAROLE | 175 |
| 10 | ACTIVITES 1 ET 2 : RESULTATS CROISES | 187 |
| 10.1 | PERIODE DE REFERENCE SUPRAMODALE..... | 187 |
| 10.2 | LIENS ENTRE LES MECANISMES DE TRAITEMENT DU RYTHME ET DU LANGAGE | 191 |
| V. DISCUSSION ET PERSPECTIVES..... | | 195 |
| BIBLIOGRAPHIE | | 215 |
| ANNEXES – PARTIE 1 | | 235 |
| ANNEXE 1 | NOTICE D'INFORMATION ET CONSENTEMENT ECLAIRE | 236 |
| ANNEXE 2 | QUESTIONNAIRE SOCIO-DEMOGRAPHIQUE | 239 |
| ANNEXE 3 | INFORMATIONS SOCIO-DEMOGRAPHIQUES DES PARTICIPANTS..... | 241 |
| ANNEXE 4 | LISTES PSEUDO-ALEATOIRES DE L'ACTIVITE 2 | 243 |
| ANNEXE 5 | MATERIEL EXPERIMENTAL DE L'ACTIVITE 2..... | 244 |
| ANNEXES – PARTIE 2 | | 245 |

LISTE DES TABLEAUX

| | | |
|---------------------|--|-----|
| Tableau 1 : | Illustration de la congruence des couples amorce-cible..... | 93 |
| Tableau 2 : | Tableau récapitulatif des 12 conditions expérimentales de l'Activité 2..... | 94 |
| Tableau 3 : | Illustration des quatre patrons de groupes métriques pour les amorces L et NL..... | 96 |
| Tableau 4 : | Récapitulatif du profil socio-démographique des participants..... | 112 |
| Tableau 5 : | Exemple de 4 frappes extraites de la SMS au tempo 600 ms..... | 116 |
| Tableau 6 : | Valeurs descriptives du nombre d'erreurs de lecture des 41 participants | 121 |
| Tableau 7 : | Indices de segmentation pour recueillir le Temps d'Initiation de la parole (TI). | 123 |
| Tableau 8 : | Valeurs descriptives des TMS moyens des participants | 136 |
| Tableau 9 : | Valeurs descriptives des stabilités moyennes du TMS..... | 141 |
| Tableau 10 : | Tableau de fréquences de la significativité des longueurs R de la SMS | 147 |
| Tableau 11 : | Valeurs descriptives de la stabilité moyenne de la SMS | 148 |
| Tableau 12 : | Valeurs descriptives de la précision de la SMS des participants | 157 |
| Tableau 13 : | Tableau de fréquences du signe de θ (moins/plus) en fonction du tempo..... | 162 |
| Tableau 14 : | Description de la stabilité et de la précision de la SMS d'un participant..... | 165 |
| Tableau 15 : | Tableau de fréquences des tempi les plus stables lors de la SMS | 165 |
| Tableau 16 : | Tableau de fréquences des tempi les plus précis de la SMS | 166 |
| Tableau 17 : | Tableau de fréquences des tempi les plus stables lors de la SMS chez les musiciens..... | 166 |
| Tableau 18 : | Tableau de fréquences des tempi les plus stables lors de la SMS chez les non-musiciens | 166 |
| Tableau 19 : | Tableau de fréquences des tempi les plus précis lors de la SMS chez les musiciens | 167 |
| Tableau 20 : | Tableau de fréquences des tempi les plus précis lors de la SMS chez les non-musiciens..... | 167 |
| Tableau 21 : | Valeurs descriptives du nombre d'erreurs de lecture moyen des participants..... | 172 |
| Tableau 22 : | Probabilités estimées d'apparition d'une erreur selon EM et TA..... | 174 |
| Tableau 23 : | Valeurs descriptives du temps d'initiation de la parole moyen des participants..... | 176 |
| Tableau 24 : | Moyennes marginales estimées du TI de la parole, selon EM et TA | 181 |
| Tableau 25 : | Tableau de fréquences des vecteurs circulaires non significatifs au test de Rayleigh..... | 188 |
| Tableau 26 : | Tableau de fréquences de l'asynchronie positive de la SMS, en fonction du tempo..... | 188 |
| Tableau 27 : | Valeurs descriptives du TMS moyen en fonction du tempo le plus stable de la SMS | 189 |
| Tableau 28 : | Valeurs descriptives du TMS moyen en fonction du tempo le plus précis de la SMS. | 189 |
| Tableau 29 : | Matrice de corrélation entre les mesures de l'Activité 1 et l'Activité 2..... | 193 |
| Tableau 30 : | Description socio-démographique (âge, années pratique musique et danse)..... | 241 |
| Tableau 31 : | Description socio-démographique (sexe)..... | 241 |
| Tableau 32 : | Description socio-démographique (musique) | 241 |
| Tableau 33 : | Description socio-démographique (instrument)..... | 242 |
| Tableau 34 : | Description socio-démographique (danse)..... | 242 |
| Tableau 35 : | Description socio-démographique (jeux vidéo) | 242 |
| Tableau 36 : | Listes pseudo-aléatoires de passation (Activité 2)..... | 243 |
| Tableau 37 : | Syntagmes (amorces et cibles) utilisés dans l'Activité 2..... | 244 |

LISTE DES FIGURES

| | | |
|--------------------|--|-----|
| Figure 1 : | Exemple de structure prosodique | 27 |
| Figure 2 : | Terminologie utilisée dans les approches de la hiérarchie des domaines prosodiques..... | 28 |
| Figure 3 : | Représentation de la structure métrique de l'énoncé « Ils dévorent leur goûter »..... | 30 |
| Figure 4 : | Illustration de l'arc accentuel..... | 32 |
| Figure 5 : | Caractéristiques phonétiques de AI et AF en français..... | 33 |
| Figure 6 : | Modèle simplifié de la reconnaissance des mots de Bagou et Frauenfelder (2002)..... | 39 |
| Figure 7 : | Modèle neurocognitif de la compréhension du langage de Friederici (2017)..... | 40 |
| Figure 8 : | Modèle à double voie ventrale/dorsale de Hickok et Poeppel (2007)..... | 42 |
| Figure 9 : | Modèle d'encodage lexical pour la production de parole d'Indefrey et Levelt (2004)..... | 43 |
| Figure 10 : | Modèle DIVA d'articulation de la parole de Tourville et Guenther (2011)..... | 46 |
| Figure 11 : | Illustration des connexions formant le faisceau frontal aslant | 47 |
| Figure 12 : | Transformée de Fourier de l'enveloppe d'amplitude d'un signal de parole | 58 |
| Figure 13 : | Illustration des conditions de congruence métrique de Cason, Astésano et al. (2015)..... | 76 |
| Figure 14 : | Tempo Moteur Spontané..... | 86 |
| Figure 15 : | Synchronisation Sensori-Motrice..... | 86 |
| Figure 16 : | Amorçage rythmique de la parole (lecture à haute voix) | 90 |
| Figure 17 : | Variation de la phase d'amorce selon l'Engagement Moteur (EM) du participant..... | 91 |
| Figure 18 : | Variation de la phase d'amorce selon le Type d'Amorce..... | 92 |
| Figure 19 : | Variation d'amorçage selon la Congruence Métrique | 92 |
| Figure 20 : | Tâche d'amorçage avec les trois Variables Indépendantes..... | 93 |
| Figure 21 : | Représentation en grille métrique de l'énoncé « Ils dévorent leur goûter » | 96 |
| Figure 22 : | Visualisation d'une amorce de type musical dans le logiciel Audacity..... | 99 |
| Figure 23 : | Illustration des six blocs expérimentaux de l'Activité 2 | 101 |
| Figure 24 : | Salle de passation expérimentale | 105 |
| Figure 25 : | Estimation du nombre de participants par le logiciel G*Power..... | 111 |
| Figure 26 : | Représentation circulaire de 4 frappes pour un extrait de la tâche de SMS | 115 |
| Figure 27 : | Illustration de la mesure du temps d'initiation (TI) de la parole | 119 |
| Figure 28 : | Représentation graphique du nombre d'erreurs de lecture des participants | 121 |
| Figure 29 : | Exemple d'une segmentation en syllabes du syntagme « L'impopulaire bûcher »..... | 123 |
| Figure 30 : | Représentation de la distribution des TMS moyens des participants | 135 |
| Figure 31 : | Evolution du TMS moyen en fonction de l'âge..... | 136 |
| Figure 32 : | Evolution du TMS moyen en fonction du nombre d'années de pratique de la danse | 137 |
| Figure 33 : | Evolution du TMS moyen en fonction de l'âge..... | 138 |
| Figure 34 : | Représentation de la répartition des IBI et du vecteur du participant Z37..... | 140 |
| Figure 35 : | Représentation graphique de la stabilité du TMS des participants..... | 141 |
| Figure 36 : | Evolution de la stabilité du TMS en fonction de l'âge | 142 |
| Figure 37 : | Evolution de la stabilité du TMS en fonction du nombre d'années de musique. | 142 |

| | | |
|--------------------|--|-----|
| Figure 38 : | Evolution de la stabilité du TMS selon le nombre d'années de pratique de la danse..... | 143 |
| Figure 39 : | Représentation graphique de la stabilité moyenne de la SMS des participants..... | 148 |
| Figure 40 : | Evolution de la stabilité de la SMS en fonction de l'âge | 149 |
| Figure 41 : | Evolution de la stabilité de la SMS en fonction du nombre d'années de musique..... | 150 |
| Figure 42 : | Evolution de la stabilité de la SMS en fonction du nombre d'années de pratique de la danse | 150 |
| Figure 43 : | Comparaison de la stabilité de la SMS en fonction du tempo (1 ^{er} modèle) | 152 |
| Figure 44 : | Comparaison de la stabilité de la SMS en fonction du tempo (2 ^{ème} modèle)..... | 153 |
| Figure 45 : | Evolution de la stabilité de la SMS selon les tempi et le nombre d'années de musique | 153 |
| Figure 46 : | Représentation graphique de la précision de la SMS des participants | 157 |
| Figure 47 : | Evolution de la précision de la SMS en fonction de l'âge | 158 |
| Figure 48 : | Evolution de la précision de la SMS en fonction du nombre d'années de musique..... | 159 |
| Figure 49 : | Evolution de la précision de la SMS en fonction du nombre d'années de danse | 159 |
| Figure 50 : | Evolution de la précision de la SMS en fonction du nombre d'années de musique..... | 161 |
| Figure 51 : | Représentation graphique du nombre d'erreurs de lecture moyen des participants | 171 |
| Figure 52 : | Evolution du nombre d'erreurs de lecture en fonction de l'âge | 172 |
| Figure 53 : | Représentation graphique du TI de la parole moyen des participants | 175 |
| Figure 54 : | Evolution du TI de la parole en fonction de l'âge..... | 176 |
| Figure 55 : | Evolution du TI de la parole en fonction du nombre d'années de musique..... | 177 |
| Figure 56 : | Comparaison du TI de la parole selon l'Engagement Moteur des participants | 178 |
| Figure 57 : | Comparaison du TI de la parole en fonction du Type d'Amorce | 179 |
| Figure 58 : | Comparaison du TI de la parole en fonction de EM et TA..... | 180 |
| Figure 59 : | Comparaison du TI de la parole selon l'Engagement Moteur et la pratique de la musique | 182 |
| Figure 60 : | Comparaison du TI de la parole en fonction de CM et du groupe métrique de l'amorce | 184 |
| Figure 61 : | Evolution du TI de la parole moyen en fonction du TMS moyen..... | 190 |

INTRODUCTION

Vous êtes-vous déjà surpris en train de bouger en rythme ou de chanter en chœur sur une musique ? D'où nous vient cette irrésistible envie d'associer des gestes avec une suite de sons entendus ? C'est la perception d'une pulsation rythmique. En effet, le rythme est cette ossature qu'on retrouve au cœur de la musique, de la parole ou encore de la danse, sur laquelle les notes, les mots et les gestes se posent, s'organisent et se déroulent dans le temps. La perception rythmique est un processus actif qui concerne à la fois le caractère d'un phénomène extérieur capté par nos sens mais également un crible attentionnel qui façonne notre perception de l'environnement. Afin d'être perçu, le rythme d'un phénomène externe doit donc fournir des repères temporels pertinents aux rythmes attentionnels pour que ceux-ci s'y synchronisent et nous permettent de décoder le signal sonore qui nous parvient.

Qu'en est-il pour la parole ? Le rythme dans la parole nous fournit de précieux indices pour découper le flux auditif. En nous synchronisant sur le rythme de la parole d'un locuteur, nous pouvons ainsi identifier les segments structurellement saillants (les accents rythmiques) qui vont participer à la constitution de groupements pertinents (les mots, les phrases, les tours de parole). Lorsque le rythme du flux de parole d'un locuteur ou les capacités de synchronisation d'un auditeur sont affectés, la communication peut également être impactée.

Certains protocoles cliniques de rééducation du rythme de la parole (aphasie non fluente, maladie de Parkinson ou encore le bégaiement) utilisent ainsi le rythme comme outil pour guider la voix. Cependant, dans les recherches s'intéressant à l'impact du rythme sur la parole tant auprès de populations cliniques que dans la population générale (Cason, Astésano, et al., 2015 ; Contreras Roa et al., 2022 ; Gould et al., 2016 ; Norton et al., 2009), les différentes modalités rythmiques employées (langage, musique, battement de la main) ne sont généralement pas comparées entre elles, ce qui permettrait pourtant de proposer des méthodes de rééducation de la parole plus efficaces. C'est donc ce que nous avons investigué dans ce travail de recherche, par la mise en place d'un protocole expérimental dont l'objet central est d'observer l'effet différentiel de variations d'amorçage rythmique sur la production de la parole chez des adultes tout-venant. L'objectif principal de notre étude est de voir dans quelle mesure la parole peut se synchroniser à différentes modalités de rythme et surtout d'identifier laquelle ou lesquelles de ces modalités sont les plus efficaces pour faciliter la production de la parole.

Dans la première partie, nous commençons par interroger le concept de rythme d'un point de vue philosophique et linguistique, afin de préciser ce que nous entendons par « une parole rythmique ». Dans un second temps, nous situons le rythme cette fois-ci selon une perspective neuropsycholinguistique pour appréhender son rôle déterminant en perception et production de la parole (cf. partie I., p. 6).

Nous poursuivons par la présentation de notre problématique et des hypothèses de recherche. Nous supposons que les formes d'amorçage rythmique suivantes ont un effet facilitant sur la production de la parole : (1) amorçage rythmique auditif avec engagement moteur par rapport à un amorçage rythmique auditif seul, (2) amorçage rythmique vocal langagier par rapport à un amorçage rythmique vocal non langagier ou musical, (3) amorçage rythmique métriquement congruent par rapport à un amorçage incongruent. Nous émettons également deux autres hypothèses concernant les liens entre les capacités de traitement du rythme et du langage, l'une à propos de l'existence d'une période de référence supramodale et l'autre à propos de mécanismes de traitement partagés pour le traitement du rythme et du langage (cf. partie II., p. 74).

Nous décrivons ensuite la méthodologie expérimentale employée. Nous avons construit deux activités expérimentales dans lesquelles nous avons recueilli des mesures relatives au Tempo Moteur Spontané et à la Synchronisation Sensori-Motrice sur différents tempi de métronome (225, 600 et 900 ms) (Activité 1 : mesure des capacités de traitement du rythme) ainsi que des mesures relatives à la précision de la lecture (nombre d'erreurs produites) et au Temps d'Initiation de la parole dans une tâche de lecture de syntagmes à haute voix après différents amorçages rythmiques (amorçage auditif avec ou sans engagement moteur du participant ; amorçage de type vocal langagier, non langagier ou musical ; amorçage métriquement congruent ou incongruent) (Activité 2 : lecture à haute voix après amorçage rythmique). Le recueil a été effectué auprès d'un échantillon de 41 participants (cf. partie III., p. 84).

La partie suivante est consacrée à l'analyse des résultats dans les différentes tâches expérimentales ainsi qu'entre les tâches (partie IV., p. 130), qui nous amènent à une discussion de nos hypothèses initiales (partie V., p. 195).

Deux types d'annexes accompagnent ce document, des annexes générales (ANNEXES – PARTIE 1, p. 235) et des annexes relatives aux analyses statistiques inférentielles, dans un deuxième volume (ANNEXES – PARTIE 2).

I. PARTIE THEORIQUE

1 UNE PAROLE RYTHMIQUE

Dans ce premier chapitre, nous nous intéresserons au caractère rythmique de la parole, en nous demandant dans quelle mesure la parole est rythmique et quelles sont les manifestations du rythme dans la parole. Nous commencerons ce cheminement à partir de considérations d'ordre phénoménologique afin d'identifier les caractéristiques du rythme qui traversent les domaines et participent à la construction du seuil de rythmicité (cf. 1.1). Nous poursuivrons notre propos pour définir plus spécifiquement la parole en tant que séquence auditive complexe, dans les pas de l'approche de la phonologie métrique (cf. 1.2, p. 13). Nous serons alors en mesure d'identifier deux fonctions principales du rythme qui s'actualisent dans la parole par des moyens divers : la segmentation du signal de parole en groupes hiérarchiques (les niveaux de constituance) et l'extraction d'une pulsation dans le signal, elle aussi hiérarchique (les niveaux de proéminences).

1.1 QU'EST-CE QUE LE RYTHME ?

Avant d'aborder la question du rythme dans la parole, nous commençons par définir ce que nous entendons par rythme, au moyen de repères terminologiques sur l'origine de ce concept (cf. 1.1.1) et à travers les différents domaines d'actualisation du rythme (cf. 1.1.2, p. 7), qui mènent à la proposition de trois critères participant à la définition du caractère rythmique : structure, périodicité et mouvement (cf. 1.1.3, p. 8).

1.1.1 REPERES TERMINOLOGIQUES

« Rien n'a été moins « naturel » que cette élaboration lente, par l'effort des penseurs, d'une notion qui nous semble si nécessairement inhérente aux formes articulées du mouvement que nous avons peine à croire qu'on n'en ait pas pris conscience dès l'origine. » (Benveniste, 1966, p. 335)

Les philosophes ioniens ainsi que les poètes lyriques et tragiques de la Grèce Antique utilisent le terme « *ρυθμος* » (« *rhuthmos* ») pour désigner une forme ou disposition particulière, « dans l'instant qu'elle est assumée par ce qui est mouvant, mobile, fluide, la forme de ce qui n'a pas de consistance organique » (Benveniste, 1966, p. 333). *Ρυθμος* peut ainsi désigner la

forme des lettres de l'alphabet chez Hérodote, les formes de l'humeur ou du caractère chez Anacréon, ou encore la forme d'un vêtement chez Euripide (Benveniste, 1966, p. 328). Par son suffixe « *-thmos* », *ρῦθμος* renvoie à une modalité particulière de forme fluide, « telle qu'elle se présente aux yeux » (Benveniste, 1966, p. 332). Dans son étymologie même, *ρῦθμος* est donc intimement lié à la perception d'une forme que prend un phénomène.

Le sens du mot rythme se spécialise ensuite avec Platon¹ pour s'appliquer à la forme des mouvements du corps humain et à la disposition des figures résultant de ces mouvements, ordonnés dans la durée. Le rythme se retrouve dès lors dans la musique, la danse, la diction ou encore le travail. Dans ce nouveau sens que prend *ρῦθμος* proposé par Platon, la forme se trouve désormais déterminée par le mètre, le « *μετρον* » (« *metron* »), « séquence ordonnée de mouvements lents ou rapides » (Benveniste, 1966, p. 334-335). Cette distinction entre rythme et mètre est plus tard reprise et discutée par Saint Augustin dans son *Traité de la Musique*² (Livre 3) : le concept de rythme contient l'idée d'un enchaînement régulier de pieds ; tandis que le mètre, en plus de se fonder dans cette définition du rythme, y ajoute la présence d'une terminaison saillante, limitant le nombre de pieds et leur adjoignant des rapports de durée. « C'est pourquoi tout mètre est rythme ; mais tout rythme n'est pas mètre » (Augustin, 1878³).

1.1.2 DOMAINES D'ACTUALISATION DU RYTHME

De nos jours, qu'il soit perçu ou agi, de manière consciente ou non, le mot rythme est utilisé pour qualifier toutes sortes de domaines : nature, corps célestes, corps humain, langue, art, société, histoire, etc. Sauvanet (2000) suggère ainsi de distinguer trois classes de domaines rythmiques : les rythmes cosmologiques, les rythmes biologiques et les rythmes anthropologiques. Les rythmes cosmologiques renvoient aux rythmes des saisons, des astres, des marées, à l'alternance du jour et de la nuit, etc., tandis que les deux autres classes se retrouvent à l'échelle de l'homme.

¹ Par exemple, dans Lois : « Cet ordre dans le mouvement a précisément reçu le nom de rythme [*ρῦθμος*], tandis qu'on appelle harmonie l'ordre de la voix où l'aigu et le grave se fondent, et que l'union des deux se nomme art choral » (Benveniste, 1966, p. 334).

² Dans l'Avertissement qui précède le Livre 1, le lecteur est informé de l'objet de la musique : « déterminer les durées successives qui divisent un mouvement et le rapport qui les ordonne entre elles », le mouvement étant à entendre dans un sens général (par ex., danse, chant).

³ Texte original écrit entre 387 et 389

Les rythmes biologiques sont des systèmes oscillants innés, caractérisés par une période, une fréquence, des phases, une acrophase⁴, une amplitude (Fraisse, 1974, p. 16). Ils peuvent être spontanés ou bien déclenchés et/ou influencés par un mécanisme externe auquel ils vont ou non se synchroniser (mécanisme ordonnateur ou synchroniseur) (Fraisse, 1974, p. 37). Fraisse (1974) distingue ainsi : les rythmes biologiques ‘spontanés’, de périodicité autonome et au déclenchement spontané (par exemple : les battements du cœur) ; les rythmes ‘déclenchés’, engendrés par une excitation non périodique (par exemple : les mouvements oscillatoires des feuilles de haricot qui surviennent lors de l’obscurité) ; les rythmes ‘induits’, dont le caractère spontané est influencé par un autre rythme avant de retrouver sa fréquence d’origine (par exemple : la synchronisation des ondes cérébrales α sur la fréquence d’un rythme lumineux externe) ; les rythmes ‘acquis’ ou ‘conditionnés au temps’, dont la périodicité est engendrée par un rythme extérieur (par exemple : la sensation de faim que l’on ressent chaque jour à la même heure à l’approche du déjeuner). Ces rythmes biologiques se retrouvent à tous les niveaux dans le corps de l’homme, de la cellule à l’activité de l’organisme dans son entier. Ils se différencient encore par leur fréquence : les rythmes rapides, de fréquence inférieure à deux heures (rythme de l’encéphale, cardiaque, respiratoire) ; les rythmes circadiens, de fréquence environ égale à vingt-quatre heures, pouvant être liés à l’alternance jour/nuit (activité/veille, variations de la température corporelle) ; les rythmes lents, de fréquence supérieure à vingt-quatre heures, pouvant être liés aux rythmes des saisons par exemple.

Les rythmes anthropologiques renvoient aux rythmes de l’activité humaine qui participent à construire la notion de communauté, c’est-à-dire aux rythmes historiques, techniques ou esthétiques (Sauvanet, 2000).

1.1.3 CRITERES DEFINITOIRES : STRUCTURE, PERIODICITE ET MOUVEMENT

La difficulté réside alors dans l’approche définitoire du rythme : chaque domaine d’application propose déjà de nombreuses définitions, spécifiques, et qui se transposent difficilement voire nullement de domaine à domaine ou d’auteur à auteur d’un même domaine. Voici par exemple quelques définitions présentées dans Sauvanet (2000) pour illustrer les multiples sens du rythme et qui paraissent opposées sur certains aspects :

⁴ L’acrophase correspond à l’emplacement du pic d’intensité/amplitude, le sommet de la fonction sinusoïdale.

1. « *Le rythme est dans n'importe quel laps de temps (non-structure). Absence d'ordre. C'est enfin le printemps – pas seulement dans l'air. Prenez comme exemple de rythme tout ce qui semble incongru (irrelevant)* », citant J. Cage (1967) (compositeur).

2. « *Le rythme est l'impression que l'on éprouve d'une régularité dans le retour des temps marqués.* », citant M. Grammont (1965) (phonéticien).

3. « *Le rythme est en même temps l'ordre, la mesure dans le mouvement et la manière personnelle d'exécuter ce mouvement.* », citant E. Jaques-Dalcroze (1920) (pédagogue de la musique).

Pour Sauvanet, une des difficultés de définition du rythme vient du fait que le rythme n'a pas d'essence propre : ce qu'ont en commun les différentes formes de rythme serait donc bien leur caractère rythmique. Sauvanet propose alors de définir le phénomène rythmique au moyen d'une association de trois critères : la structure, la périodicité et le mouvement. Chacun des trois critères doit être associé à un ou à deux autres critères pour qu'un phénomène devienne rythmique, mais la présence d'un seul critère ne peut suffire pour atteindre le seuil de rythmicité.

Ainsi, Sauvanet relie le critère « S comme Structure » au *skhêma* grec, que Benveniste (1966) rapprochait du sens initial de *ρυθμος* dans certains de ses emplois. Le terme de structure (plutôt que forme) choisi par Sauvanet renforce l'idée de construction, d'organisation, de « tout constitué par l'ensemble des relations entre ses parties » (2000, p. 167). Le critère de structure comporte deux paramètres quantitatifs et deux paramètres qualitatifs : le rapport, le poids, le grain et le ton, dont l'expression varie selon la modalité sensorielle convoquée. En prenant l'exemple d'un objet sonore perçu (musique ou parole), la structure rythmique « S » s'y manifeste par la variation des paramètres de durée (rapport), d'intensité (poids), de timbre (grain) et de hauteur (ton). Une note ou syllabe peut donc être perçue comme accentuée lorsque la modification d'un ou plusieurs des quatre paramètres combinés la fait ressortir en tant que forme structurée par rapport au fond inaccentué. Sauvanet insiste enfin sur le « seuil de structuration », c'est-à-dire la structure minimale requise pour pouvoir parler de phénomène rythmique : celle-ci renverrait à l'existence d'un contraste, d'une alternance entre deux éléments séparés par un intervalle.

D'autre part, Sauvanet associe « P comme Périodicité » au *periodos* grec. Le critère de périodicité renvoie aux phénomènes perçus comme des cycles, des répétitions. Ils impliquent la présence d'un repère qui revient à intervalles réguliers et fait office de pulsation. La

périodicité fait exister un phénomène dans un « présent épais »⁵ spatial et/ou temporel, tiré entre un avant et un après, l'après étant attendu sur la base de l'avant et du présent. Les trois paramètres quantitatifs qui composent la périodicité sont sa période, sa fréquence et son nombre d'apparitions. Un motif rythmique musical pourra ainsi avoir par exemple une période de 0.5 secondes, une fréquence correspondante de 2 Hz (qui renseignera sur le tempo), et apparaître (se répéter) cinq fois dans le morceau musical. De même qu'il existe un seuil de structuration, le seuil de périodicité se situe au niveau du troisième repère : nous retrouvons cette notion de seuil de périodicité chez Ceriani (2000, p. 105), pour qui « il n'y a du rythme qu'à partir d'un troisième événement, succédant au premier et à sa répétition, lequel fonde l'attente du perceuteur ».

Pour finir, Sauvanet associe « M comme Mouvement » au *metabolè* grec. Le mouvement, c'est la vie, la transformation, l'inattendu, ce qui déborde du champ d'action de la structure ou de la périodicité. Lors de l'écoute d'un extrait musical, l'introduction du critère de mouvement (par exemple par des retards) va créer un effet de surprise en déjouant les attentes de l'auditeur, faisant naître une émotion et donc vivre le morceau de musique. Le mouvement crée une différence, il fait naître le rythme, et par là-même, il fait preuve « d'auto-genèse par auto-différenciation » (Sauvanet, 2000, p. 191).

Selon cette conception, un phénomène ne présentant qu'un seul des trois critères ne pouvant pas être considéré comme rythmique, quatre combinaisons entre les trois critères sont possibles et renvoient ainsi au phénomène rythmique : les combinaisons Structure-Périodicité (SP) ; Mouvement-Structure (MS) ; Périodicité-Mouvement (PM) et Structure-Périodicité-Mouvement (SPM).

Ce faisant, l'association de la structure à la périodicité (sans mouvement) (SP) fait émerger l'aspect métrique répétitif du rythme. Nous retrouvons ici la notion de mètre évoquée plus haut comme une séquence ordonnée présentant une terminaison saillante, que Sauvanet rapproche du terme de « mesure » de S. Arom (1986) (« soubassement mesuré du rythme »). En devenant périodique, la structure devient un modèle prédictif de la structure à venir, à partir de la structure qui est répétée.

Par ailleurs, l'association du mouvement et de la structure (sans périodicité) (MS) mène à des phénomènes rythmiques apériodiques, de syncope. La structure se métamorphose sans

⁵ Missire, R. (2021, 12 mars). *La notion de rythme – éléments d'analyse définitionnelle* [communication orale]. Séminaire du Groupe de Travail sur le Rythme, Toulouse.

régularité, comme dans certaines compositions de musique contemporaine, par exemple le morceau *Constructions* du compositeur John Cage auquel Sauvanet renvoie.

Aussi, la périodicité associée au mouvement (sans structure) (PM) se retrouve typiquement dans les rythmes naturels (biologiques et cosmologiques), caractérisés à la fois par leur caractère périodique mais aussi par leur nature imprévisible, vivante (mouvement).

Finalement, l'association des trois critères ensemble (SPM) renvoie au rythmique au sens fort, « à la fois discontinu (structure par intervalles) et régulier (périodicité), tout en se laissant une tolérance d'irrégularité (mouvement), et en se donnant globalement comme une continuité (l'ensemble structure-périodicité-mouvement) » (Sauvanet, 2000, p. 196). De plus, selon Astésano (2001, p. 45), la distorsion (associée ici au Mouvement) par rapport au mètre (Structure et Périodicité) est retrouvée dans la plupart des performances rythmiques des activités humaines et résulte de diverses contraintes de production, individuelles ou situationnelles, reflets de l'esthétique humaine.

Les trois critères fondamentaux de structure, périodicité et mouvement permettent ainsi de proposer différentes acceptions du rythme, selon leurs combinaisons ou leurs degrés d'actualisation, qui peuvent faire émerger des définitions plus spécifiques. Si nous revenons aux trois exemples de définitions du rythme exposées plus haut (p. 9), il apparaît donc que la première définition (J. Cage) conçoit le rythme en s'appuyant essentiellement sur le critère de mouvement qui crée la surprise, l'incongruité (M), face aux attentes induites par S et/ou P. La seconde définition (M. Grammont) insiste sur le critère de périodicité (P) des temps marqués (S), sans évoquer le mouvement. Dans la troisième définition (E. Jaques-Dalcroze) en revanche, les trois critères apparaissent équilibrés *via* les notions de « mesure » (SP) et de « manière personnelle » (M).

En résumé, Sauvanet propose donc de considérer comme rythme « tout phénomène perçu, subi ou agi, auquel un sujet peut attribuer au moins deux des critères suivants : structure, périodicité, mouvement » (Sauvanet, 2000, p. 195). Cette définition invite à une double lecture, impliquant à la fois un versant phénoménologique (relatif au sujet) et un versant terminologique ou critériologique (relatif à l'objet).

Le rythme serait ainsi une sorte de pont connectant le monde des objets, réel et émergent, au monde du sujet, phénoménal et contraignant. En tant que tel (pont, intermédiaire), Ceriani (2000) propose de rapprocher le terme 'rythme' des concepts de *Gestalt* ou de morphologie, soit des structures dynamiques qui façonnent la perception tout en se

modifiant au gré des expériences sensorielles. Pour être perçu, subi ou agi comme un phénomène rythmique, un objet doit être saillant, c'est-à-dire qu'il doit se détacher sur un fond amorphe. Lorsque la saillance est biologiquement signifiante, qu'elle s'impose, on parle alors, selon Thom, de prégnance (1988, cité par Ceriani, 2000, p. 108). Dans le cas du rythme, la saillance ou prégnance d'un objet est dépendante de son seuil de rythmicité, relatif aux combinaisons des critères de structure, périodicité et mouvement. Pour Ceriani (2000, p. 109), la prégnance d'un rythme participe à stabiliser le système perceptif, en régulant biologiquement l'activité symbolique. Sauvanet (2000, p. 179) évoque par exemple la périodicité associée aux formules utilisées dans les rites ou la magie : la répétition les rend prégnantes, imposantes et leur confère leur influence ou supposé pouvoir.

1.2 LE RYTHME DANS LA PAROLE

Dans la parole, le phénomène rythmique s’instancie à différents niveaux. Au niveau moteur, la production de parole convoque une activité motrice coordonnée entre les mouvements de la cage thoracique (coordination pneumo-phonatoire), la mobilisation des structures laryngées et des articulateurs supra-glottiques⁶ (Grabski & Sato, 2016). L’alternance cyclique des mouvements d’ouverture et fermeture de la mâchoire, associée aux syllabes parlées, mène à des changements rythmiques dans l’amplitude de la parole : le mouvement d’ouverture est associé aux segments vocaliques et à une haute amplitude, et le mouvement de fermeture aux segments consonantiques et à une basse amplitude (MacNeilage, 1998). Au niveau acoustique, les trois critères de rythmicité (SPM, voir *supra*) se retrouvent dans les variations de la matière sonore (durée, intensité, hauteur et timbre) qui discrétisent des unités ou créent des groupements structurés (S) et dans la succession temporelle quasi (M)-régulière (P) de ces unités ou groupements.

Il est admis que l’étude du rythme dans la parole relève de la prosodie, « domaine des propriétés métriques, tonales et temporelles des langues » (Di Cristo, 2010). La prosodie s’applique de manière évidente à la parole audible (produite et perçue) mais également à la parole silencieuse, comme en lecture par exemple où elle se fait le reflet d’une parole interne (Di Cristo, 2010). Dans sa fonction linguistique, la prosodie participe à l’actualisation et la structuration des autres niveaux d’organisation linguistique d’un énoncé (phonétique, phonologie, lexicale, syntaxe, sémantique, pragmatique). La prosodie exerce également deux autres fonctions : extralinguistique, en fournissant des indications sur l’identité du locuteur (par exemple son âge, son sexe, son origine régionale) et paralinguistique, à propos de l’état d’esprit du locuteur (par exemple à travers le ton employé) (Astésano, 2017, p. 19). Par ailleurs, Astésano (2017, p. 20-21) distingue les paramètres phonologiques (accentuation, intonation, rythme) et acoustiques (hauteur, durée, intensité) en analyse prosodique. Les paramètres acoustiques, que l’on peut associer au critère de structure de Sauvanet (2000), matérialisent les paramètres phonologiques en se combinant selon les lois d’une grammaire prosodique spécifique à chaque langue. Au côté de ces variations de structure prosodique, il semble cependant que certains universaux sont décrits à travers les langues, tels que les allongements touchant le mot initial (généralement sur la consonne initiale) et le mot final (généralement la

⁶ Les articulateurs actifs sont les lèvres, la langue, la mandibule, le voile du palais, l’épiglotte et la glotte ; les articulateurs passifs sont les dents, les alvéoles, le palais dur, la luvette et la paroi pharyngale (Astésano, 2014).

rime finale ou dernière syllabe accentuée) qui marquent le phrasé d'un groupe prosodique (Turk & Shattuck-Hufnagel, 2014).

Aussi, comme l'explique Di Cristo (2003) les concepts de métrique et de rythme sont parfois utilisés de manière indifférenciée dans les recherches portant sur la parole et devraient en fait être mieux différenciés. Di Cristo suggère en effet que la métrique serait une forme abstraite, organisée, plus contrainte que le rythme, qui lui, tout en découlant des contraintes métriques serait une forme de surface, dynamique car sensible aux événements. Selon la même idée, la métrique et le rythme se différencient encore en fonction du niveau cognitif de représentation auquel ils renvoient : le mètre est rattaché au niveau profond de représentation des schèmes (ou *patterns*) métriques où est spécifiée l'accentuation (syllabe forte) potentielle ; tandis que le rythme correspond au niveau de surface où se réalisent concrètement ces schèmes métriques, affectés par des contraintes propres à la réalisation de l'énoncé (contraintes lexicales, syntaxiques, sémantico-pragmatiques).

Il nous paraît important de retenir également que, dans le cadre de la phonologie métrique (cf. 1.2.2, p. 25) dont se réclament Di Cristo (2003) et Astésano (2001), le rythme (métrique) est considéré comme l'« ossature » de l'organisation prosodique, prééminent à l'intonation et système noyau régissant les relations entre proéminences. Le rythme de la parole peut ainsi être défini comme « l'organisation temporelle des proéminences qui participent de la structuration prosodique des énoncés » (Astésano, 2001, p. 9-10). Toute langue posséderait ainsi des représentations rythmiques sous-jacentes (représentations métriques) qui s'actualiseraient différemment en surface (représentations rythmiques) en fonction de la variabilité structurelle et aléatoire (Astésano, 2001, p. 27). Par exemple, la métrique traditionnelle voudrait que le pied⁷ le plus naturel soit iambique en français (groupement métrique composé d'une syllabe faible puis forte – « xX »)⁸ mais trochaïque en anglais (« Xx ») ; l'organisation du rythme reposerait donc sur la réalisation effective de ces schèmes dans la parole ordinaire (Di Cristo, 2003).

Ainsi, pour Di Cristo (2003), reprenant le modèle de Drake (1998), rythme et mètre sont deux constructions perceptives subjectives, impliquées dans les processus de traitement des séquences auditives complexes : la segmentation du signal en groupes, l'extraction d'une

⁷ Un pied métrique est constitué d'une syllabe accentuée accompagnée éventuellement d'une ou plusieurs syllabes inaccentuées (Cutler & Norris, 1988).

⁸ Nous verrons dans la partie 1.2.2.3 (p. 32) que le pied du français pourrait plutôt être spondéïque, caractérisé par une borne gauche (l'accent initial) et une borne droite (l'accent final) (X-X) (Astésano, 2017).

régularité temporelle (la pulsation) dans le signal, et l'intégration de ces deux premiers processus dans des structures hiérarchiques. Les groupes segmentés sont ainsi rassemblés en groupes d'ordre supérieur, et les battements de pulsation sont associés à différents niveaux hiérarchiques selon leur relative saillance. A partir de là, la segmentation du signal en groupements hiérarchiques dans la parole mène à la création de différents niveaux de constituance prosodique, auxquels sont associés différents niveaux hiérarchiques de prééminence. Cette organisation hiérarchique du traitement de la parole permet d'appréhender des séquences longues et complexes, par la génération d'attentes portant sur différents niveaux (syllabique, accentuel, etc.).

Comme ces niveaux de représentation hiérarchiques sont directement impliqués dans les modèles développés pour mieux comprendre les processus de traitements en perception et production de la parole, il nous paraît crucial de les aborder dès la section suivante. Ainsi, nous présenterons les processus généraux de traitement des séquences auditives complexes (segmentation en groupes et extraction de la pulsation) (cf. 1.2.1) avant de les aborder plus spécifiquement pour la parole avec l'approche de la phonologie métrique (cf. 1.2.2, p. 25).

1.2.1 TRAITEMENT DES SEQUENCES AUDITIVES RYTHMIQUES COMPLEXES

Nous présentons tout d'abord les limites sensorimotrices qui contraignent la perception et la production de séquences rythmiques (cf. 1.2.1.1), puis les mécanismes généraux de segmentation en groupements (cf. 1.2.1.2, p. 17) et d'extraction de la pulsation (cf. 1.2.1.3, p. 19).

1.2.1.1 CONTRAINTES SENSORIMOTRICES

Deux échelles de temps sont convoquées en perception auditive : d'une part, un régime temporel, où les événements acoustiques se succèdent de manière isolée et peuvent se manifester rythmiquement, et d'autre part un régime spectral, où les événements se succèdent trop rapidement pour être perçus de manière isolée, ce qui implique que ces événements sont plutôt perçus comme un son continu caractérisé par ses composantes spectrales (mélodie, timbre) (A. Albert & Grice, sous presse). Suivant cette double temporalité perceptive, un stimulus auditif peut ainsi être perçu comme un phénomène rythmique si la période de ses intervalles (sa vitesse de battement) est située entre 15-20 cs et 150-200 cs⁹ (Fraisse, 1974, p.

⁹ 15-20 cs = 150-200 ms = 5-6,7 Hz ; 150-200 cs = 1500-2000 ms = 0,5-0,67 Hz.

78) : si sa fréquence est plus basse (donc que les battements sont trop éloignés), les battements ne seront plus perçus comme faisant partie d'un même groupe rythmique ; en revanche, si sa fréquence est plus élevée (donc que les battements sont trop rapprochés), le son sera perçu comme une continuité (par exemple : une note chantée à une fréquence de 415 Hz¹⁰ sera perçue comme une hauteur et non comme une série de battements). On retrouve également ce phénomène avec les images stroboscopiques du cinéma : un enchaînement de 24 images par seconde (soit des battements très rapprochés) est perçu comme une image animée, continue, et non comme une saccade d'images (Sauvanet, 2000, p. 180-181).

En ce qui concerne le versant de la production rythmique, Piaget (1942) indique que le rythme est la forme la plus élémentaire de l'action, et ce dès les premières semaines de vie : les mouvements impulsifs ou les réflexes sont constitués par l'alternance d'une phase ascendante/agoniste (ouvrir la bouche par exemple) et d'une phase descendante/antagoniste (fermer la bouche, par exemple). Ces mouvements régulés, plus tard volontaires, sont soumis aux limites biomécaniques du corps humain : en ce qui concerne par exemple la battue du doigt (*finger tapping*), un sujet pourra réussir à produire au maximum jusqu'à 5-7 battements par seconde (5-7 Hz), soit des intervalles de 150-200 ms (Repp, 2005).

Finalement, la perception et la production rythmiques sont aussi contraintes par la culture du sujet : les formes rythmiques sont conditionnées par des habitudes culturelles et sociales, qui amènent les sujets à grouper différemment l'information symbolique (Ceriani, 2000, p. 33). Sauvanet (2000, p. 79-80) nous donne l'exemple de la segmentation d'un signal sonore, qu'il emprunte à l'ethnomusicologue Jean During (1984) : « Si un Indien, un Arabe et un Européen écoutent par exemple le bruit du train sur les rails, chacun d'eux organise ce son d'une autre manière ». Ainsi, le rythme est commun à des individus au sein d'un même groupe socio-culturel et opère comme une modalité culturellement déterminée structurant les phénomènes, que l'on retrouve par exemple dans la parole, les gestes ou la musique. Le rythme participe ainsi à homogénéiser le groupe social (Sauvanet, 2000, p. 43).

En résumé, l'objet rythmique perçu ou produit est donc contraint par des automatismes formels de pensée dynamiques en interaction avec les contraintes sensorimotrices et culturelles de l'individu, qui sont de plus, déterminés par l'environnement. En ce sens, le rythme est une structure dynamique ancrée aussi bien dans le corps (sensorimoteur et cognitif) du sujet que dans son rapport au monde collectif et son environnement culturel.

¹⁰ 415 Hz = 2,4 ms.

1.2.1.2 SEGMENTATION EN GROUPEMENTS

Lorsque les contraintes sensorimotrices sont opérationnelles, les objets émergeant à partir du monde réel peuvent être captés par les sens d'un individu. Ils peuvent ensuite être *perçus* en tant que phénomènes rythmiques grâce à la propension du sujet à regrouper naturellement des éléments isolés (Ceriani, 2000, p. 101). Il existerait donc deux stades perceptifs, correspondant à la perception des qualités sensibles et des qualités formelles ; les qualités sensibles se rapportant à la perception par les sens (par exemple des stimuli sonores isolés) et les qualités formelles se rapportant à la perception d'une structure relationnelle entre les objets captés par les sens (par exemple une mélodie regroupant en un ensemble ces stimuli sonores) (d'après Ehrenfels (1890)¹¹, cité par Ceriani, 2000, p. 119).

Le groupement rythmique qui caractérise les qualités formelles n'existe pas en tant que tel dans la nature. Il découle d'un besoin économique d'organisation de l'esprit humain, de manière à assimiler ou dissimiler les formes perçues. Ceriani y fait référence par l'expression de « geste perceptif fondamental d'organisation du contenu » (Ceriani, 2000, p. 153). Rythme et groupement sont d'ailleurs considérés par Piaget (1942) comme des « structures fondamentales de la vie psychique »¹². Cette tendance au regroupement est également le principe élémentaire de la Théorie de la Forme (*Gestalt*)¹³, laquelle énonce des lois d'organisation¹⁴ qui prédisposent l'organisme du sujet dans ses interactions avec l'environnement : les contraintes formelles qui découlent de ces lois agissent à la fois sur le versant du décodage et de l'encodage et participent à dégager ou générer des formes prégnantes, des catégories et donc du sens, forgeant ainsi des automatismes de traitement (Ceriani, 2000, p. 56). On retrouve alors ici l'idée d'un « rythme forme » dynamique déjà inhérente au *ρυθμος* grec (Benveniste, 1966) (cf. p. 6).

Selon Fraisse (1974), le groupement rythmique peut se faire de manière objective (selon la saillance de l'objet) ou subjective. En intégrant et liant des formes successives spatio-temporelles en un tout simultané, le groupe crée un « présent psychologique », relatif à la taille du groupement (Fraisse, 1974). Lors de l'écoute subjective, les stimulations semblent se

¹¹ Dans son essai *Gestaltqualitäten* qui introduit la notion de forme (*Gestalt*).

¹² Piaget introduit également la structure de « régulation », qui permet d'atteindre un équilibre en résolvant les alternances engendrées par le rythme, et qui devient groupement dès lors qu'on lui applique des opérations de la pensée logique (Piaget, 1942).

¹³ La Théorie de la Forme (ou *Gestalttheorie*) naît en Allemagne au début du 20^{ème} siècle (Ceriani, 2000, p. 118).

¹⁴ Lois d'organisation de la Gestalt : 1/ loi de proximité ; 2/ loi de similarité ; 3/ loi de délimitation ; 4/ loi du contenant ou de la bonne forme ; 5/ loi du mouvement commun ; 6/ loi de l'expérience ; 7/ loi de la prégnance (Ceriani, 2000, p. 120).

grouper par deux, trois ou plus rarement quatre éléments séparés par un intervalle plus long faisant office de pause. Lorsque davantage d'éléments sont groupés par un moyen objectif, ils vont avoir tendance à être regroupés en sous-ensembles. En appliquant le principe de rythmisation subjective par exemple à des gouttes d'eau tombant à intervalles réguliers, le sujet va introduire une différence de durée subjective qui rendra les intervalles entre deux groupes de gouttes subjectivement plus longs et les intervalles entre gouttes d'un même groupe subjectivement plus courts. La première (ou dernière parfois) stimulation est souvent perçue plus intensément que les suivantes et suivie d'un intervalle légèrement plus long. Dans le domaine sonore, cette saillance peut être créée ou contrainte perceptivement comme une augmentation ou variation d'intensité, de durée, de hauteur ou de timbre, sur la base de paramètres structurels stables.

Selon ce principe décrit par Fraisse (1974), nos capacités perceptives nous permettent d'individualiser deux à trois durées au sein d'un groupe rythmique. Les stimuli relevant de durées voisines sont donc ensuite assimilés ou dissimilés afin de rejoindre ces deux ou trois classes de durées (Fraisse, 1974, p. 406). Toujours selon lui, certaines formes rythmiques prégnantes dominant dans la production spontanée : les temps courts et les temps longs, liés par un rapport de proportion – les temps courts étant deux fois plus courts que les temps longs, et le rapport d'égalité étant le plus fréquent dans une suite d'intervalles. C'est au tempo moteur spontané¹⁵ de l'individu que cette différence entre temps courts et temps longs est la plus juste. Le tempo spontané correspond ainsi à la forme rythmique la plus prégnante (Fraisse, 1974, p. 100).

En ce qui concerne plus spécifiquement le signal de parole, la segmentation en groupements est déterminée par l'alternance entre syllabes accentuées et inaccentuées : les syllabes accentuées se détachent perceptivement du fond sonore constitué de syllabes inaccentuées et facilitent ainsi l'accès lexical par rapport à une segmentation basée sur les représentations phonétiques segmentales ou syllabiques (Cutler & Norris, 1988). Il est admis que la nature de la prééminence perçue va conditionner sa place dans le groupement : un accent actualisé par une augmentation d'intensité ou de hauteur va être perçu comme débutant un groupement, tandis qu'un accent actualisé par un allongement de durée sera perçu comme terminant le groupement (A. Albert & Grice, sous presse ; Fraisse, 1974).

¹⁵ Le tempo moteur spontané correspond au tempo auquel l'individu est le plus à l'aise pour frapper des temps réguliers avec sa main sur une surface. Cette notion sera plus amplement détaillée en partie 2.3.1 (p. 61).

1.2.1.3 EXTRACTION D'UNE REGULARITE

Selon Drake (2010), l'extraction d'une régularité est complémentaire à la segmentation en groupements en ce qu'elle permet de créer une continuité et donc une relation entre les groupements. L'extraction d'une régularité permet ainsi à l'auditeur de former des attentes sur la suite des événements sonores en fonction du début de la séquence.

Pour Lenc et al. (2021), la représentation interne d'une pulsation n'est pas mesurable directement et nécessite de passer par la mesure d'un processus secondaire (corrélat neuronal ou comportemental) répondant à trois caractéristiques : le processus secondaire doit (1) se dérouler dans le temps, (2) pouvoir être mesuré de manière précise et (3) être couplé de manière univoque avec les représentations internes. Les représentations internes pourraient ainsi se refléter dans les fluctuations dynamiques de l'activité neuronale ou le mouvement corporel. En effet, le mouvement corporel est un bon candidat pour illustrer les représentations rythmiques internes en raison des forts liens entre motricité et rythme (nous reviendrons sur les liens entre motricité et rythme en partie 2.2.2, p. 52). Cependant, ce prisme méthodologique impose de tenir compte des limites biomécaniques qui contraignent la réalisation de mouvements et la perception rythmique sonore, et donc limitent l'étendue des mesures rythmiques internes (cf. 1.2.1.1, p. 15).

Pour mieux comprendre les liens entre le processus d'extraction d'une régularité et sa manifestation comportementale susceptible d'être saisie à travers la temporalité, nous abordons dans les sections suivantes le cas particulier de la période de référence (ou préférence de pulsation) qui se situerait aux alentours de 600 ms (cf. 1.2.1.3.1), puis la question plus spécifique de recherche d'une pulsation dans la parole (cf. 1.2.1.3.2, p. 21), pour aboutir à l'approche de la phonologie métrique (cf. 1.2.2, p. 25).

1.2.1.3.1 PERIODE DE REFERENCE

L'extraction de la pulsation se ferait plus facilement selon une période de référence, qui varie selon les individus. Cette période de référence est associée aux fluctuations de l'attention et se traduit par un traitement optimal des événements survenant à une période proche de la période de référence. Par exemple, lors de l'écoute d'un morceau de musique, l'auditeur a la possibilité de taper du pied 'en rythme' sur différents niveaux de battement (la noire, la croche, la mesure etc.). Le niveau choisi sera celui se rapprochant le plus de sa période de référence (Drake, 2010). Drake, Jones, et al. (2000) différencient ainsi les termes de 'période' de référence et de 'niveau' de référence, la période de référence étant reflétée par le Tempo Moteur Spontané

(TMS) et le niveau de référence¹⁶ étant reflété par le tempo auquel un stimulus externe rythmique complexe est le plus susceptible d'entraîner une synchronisation de la période de référence. Les autrices présentent les résultats d'une étude dans laquelle elles comparent la période et le niveau de référence, et observent des corrélations positives significatives entre les deux variables (TMS moyen et niveau de référence), ces corrélations étant plus fortes chez les musiciens, avec toutefois l'ensemble des participants qui présentent un tempo du niveau de référence plus lent que leur TMS (période de référence).

La période de référence se situerait dans une gamme de tempi allant d'environ 380 à 800 ms, le tempo ralentissant avec l'âge avec une moyenne pour le TMS aux alentours de 400 ms dans l'enfance et 600 ms à l'âge adulte (Dalla Bella et al., 2017 ; Drake et al., 2000 ; Fraise, 1974 ; Provasi & Bobin-Bègue, 2003). Bien que la variabilité inter-individuelle soit importante, Fraise (1974) indique que ce n'est pas le cas de la variabilité intra-individuelle, le TMS restant stable sur différentes mesures. De plus, la période de référence serait plus lente chez les musiciens que les non-musiciens (Drake et al., 2000). Concernant la période de préférence de perception rythmique, Patel et Iversen (2014) rapportent une meilleure perception de la pulsation musicale pour des périodes allant de 400 ms à 1,2 s avec une préférence aux alentours de 600 ms, et une gamme de tempi pour les musiques sur lesquelles nous avons tendance à danser allant de 400 à 500 ms avec une préférence aux alentours de 451 ms. L'avancée de l'âge et de l'expertise musicale serait également liée à une amélioration du traitement hiérarchique métrique, avec une plus grande étendue de niveaux de synchronisation choisis lors de l'écoute d'un morceau musical, ainsi qu'à une amélioration de performance en synchronisation et discrimination de tempi (Drake et al., 2000).

En ce qui concerne la parole, on retrouve également une période moyenne aux alentours de 600 ms. L'intervalle inter-accentuel moyen est de 550 ms dans différentes langues (cf. Fant et al., 1991 pour le suédois, le français et l'anglais ; Fraise (1974), rapportant les travaux de Wallin (1901) sur l'anglais). Astésano (2001) relève quant à elle une durée moyenne de 600 ms pour l'intervalle inter-accentuel en français sur différents styles de parole.

Il est également intéressant de se référer aux données temporelles associées à la production de mots, syntagmes ou phrases en production de la parole. En effet, le Temps d'Initiation (TI) de la parole, qui recouvre en quelque sorte ce qu'on appelle « temps de réponse » (recouvrant temps de réaction et de latence) en psycholinguistique expérimentale,

¹⁶ Fraise (1974) parle quant à lui de 'tempo préféré'.

peut être considéré comme un des marqueurs reflétant le traitement du rythme de la parole du locuteur (Späth et al., 2016). Ce TI de la parole se situe aux alentours de 600-700 ms pour des locuteurs adultes neurotypiques, en témoignent les résultats issus de diverses études manipulant du matériel linguistique dans des tâches de production verbale, ou les modèles intégrant des signatures temporelles du traitement du langage. Par exemple, selon Gould et al. (2016), le TI de la parole est de 513 ms pour la lecture de mots et de 769 ms pour des pseudo-homophones ; Indefrey et Levelt (2004) rapportent un TI de 600 ms en production de mots sur entrée visuelle (dénomination d'image et lecture de mot) et Speer et Wilshire (2014) rapportent les résultats de deux expériences dans lesquelles ils observent un TI en dénomination d'image de 744 ms (expérience 1) et 765 ms (expérience 2) pour un mot de fréquence lexicale élevée, et 818 ms (expérience 1) et 829 ms (expérience 2) pour un mot de fréquence lexicale basse. D'autre part, l'étude de Späth et al. (2016) met en évidence un TI de 701 ms en lecture de phrases, Zhang et Zhang (2019) rapportent un TI de 679 ms en lecture de syntagmes. Les aspects psycholinguistiques en terme de traitements d'informations linguistiques impliqués en production de la parole et influençant le temps d'initiation sont déterminants, c'est pourquoi nous en exposerons les principes généraux dans la partie 2.1.2.1 (p. 43).

Pour appréhender le rythme dans la parole, on se réfère ainsi aux manifestations comportementales d'une période de référence démontrée par les données du TI, mais aussi par la recherche d'une pulsation dans la parole à travers la description de son organisation temporelle (débit, pauses, saillances accentuelles, métrique).

1.2.1.3.2 UNE PULSATION DANS LA PAROLE

La rythmicité de la parole, perçue ou produite, n'est pas strictement régulière d'un point de vue acoustique, mais elle n'est pas pour autant aléatoire : les intervalles temporels pertinents sont relativement réguliers et servent d'indices à l'auditeur et au locuteur pour prédire les événements acoustiques futurs. Sur la base de la vaste littérature et nombreux travaux publiés dans le domaine, A. Albert et Grice (sous presse), posent que l'isochronie rattachée à la parole se trouverait non pas dans le signal produit mais dans les oscillations neuronales endogènes : la quasi-périodicité des oscillations internes servirait de repère temporel face au signal de parole perçu ou à produire, afin d'associer un sens aux variations dynamiques agogiques¹⁷ du débit de parole. Ces variations s'instancient par exemple quand un locuteur opère un changement de débit au sein de son énoncé pour introduire une séquence de reformulation de son propos (une

¹⁷ Accélération ou décélération.

incise) ou pour mettre en avant un élément plus important dans son énoncé, avant de le poursuivre. Les variations peuvent encore être instanciées par des interlocuteurs pour réguler et adapter les tours de parole entre eux (un ralentissement du débit annonce la fin d'un énoncé ou d'un tour de parole).

La description linguistique du rythme de la parole et ses corrélats psycholinguistiques est l'endroit de débats encore actuels. Traditionnellement, la typologie rythmique des langues les distingue entre elles selon une rythmicité (ou *isochronie*) accentuelle, syllabique ou moraique, dont l'adéquation psycholinguistique (en termes de traitements et de capacités) a été notamment remise en cause par des travaux s'intéressant, par exemple, à l'influence prégnante de la complexité syllabique variable selon les langues sur la perception en acquisition précoce (Ramus et al., 1999). D'autre part, il semblerait que la régularité rythmique n'existe qu'au niveau perceptif, et non au niveau acoustique, et que les différents types de rythmicité peuvent coexister dans une même langue (Astésano, 2001, p. 32-33). De plus, l'accentuation n'étant pas basée sur les mêmes composantes phonétiques¹⁸ selon les langues, les locuteurs (naïfs ou phonéticiens) ont alors tendance à percevoir les saillances accentuelles d'une langue étrangère en fonction des critères d'accentuation de leur(s) propre(s) langue(s) (Arvaniti, 2009). Par exemple, Arvaniti cite Roach (1982) qui suggère que les auditeurs anglophones ne perçoivent pas de différence de proéminence en français en raison d'une absence de réduction de la richesse spectrale entre les voyelles accentuées et inaccentuées : les syllabes semblent ainsi avoir toutes la même importance. La difficulté se trouve être alors d'identifier la ou les unités rythmiques de référence : syllabe, *p-center*, proéminence, etc. ?

Plus précisément, dans le cadre de la phonologie métrique (cf. 1.2.2, p. 25), la rythmicité perçue dans la parole peut être appréhendée par l'étude de l'organisation temporelle des proéminences (Astésano, 2001). Cette organisation temporelle métrique suppose que les éléments acoustiques contrastés du signal de parole se regroupent en tenant compte des contraintes sensorimotrices du régime temporel, sans que soit convoqué le critère d'isochronie (A. Albert & Grice, sous presse). La pulsation, ou sensation de battement, est donc perçue dans l'alternance d'éléments forts (les proéminences ou accents) et faibles qui se succèdent dans le flux de parole.

¹⁸ Nous reviendrons sur les caractéristiques phonétiques de l'accentuation du français dans la partie 1.2.2.3.2, p. 33.

Quant à la place précise de la proéminence dans la syllabe, Morton et al. (1976) proposent qu'elle ne serait pas déterminée par des critères acoustiques ou physiques, mais bien perceptifs. Au cours de la conception de matériel expérimental pour des expériences sur la mémoire, Morton et al. (1976) sont confrontés à la difficulté de construire une liste de mots prononcés à intervalles réguliers, les intervalles acoustiquement réguliers étant perçus comme irréguliers. Se pose alors la question de savoir quel est cet élément qui génère la perception de pulsation ou de régularité et qu'ils décident de nommer le « *perceptual center* » ou encore « *p-center* ». Les auteurs proposent ainsi une tâche d'élicitation afin de déterminer la position de ce *p-center* dans le mot : des participants doivent manipuler des séquences de mots pré-enregistrés de telle sorte que ceux-ci soient perçus comme régulièrement espacés au sein de la séquence. Les séquences produites ne sont étonnamment pas isochrones acoustiquement. Il en ressort que la perception de régularité dans la parole est un phénomène complexe qui n'est pas lié strictement à des paramètres acoustiques tels que l'attaque du mot, l'attaque de la voyelle ou encore le pic d'intensité de la voyelle, mais correspondrait plutôt à un « moment psychologique d'occurrence ». Morton et al. transposent également le concept de *p-center* à la production orale, le « p » devenant l'initiale de « production » ou « performance » (Morton et al., 1976). Nous pouvons faire l'analogie entre ce *p-center* dans la parole et l'expression « jouer au fond du temps » en musique, qui place le battement de pulsation non pas dès l'attaque du son (par exemple un son de percussion) mais un peu plus tard, comme guidé par la recherche d'un centre de gravité du battement.

Selon le même ordre d'idées, Fowler (1979) propose d'aller plus loin dans la mise en évidence du *p-center* au moyen de cinq études investiguant les versants de la perception et de la production de parole isochrone. Elle montre ainsi que lorsqu'un locuteur a pour consigne de prononcer des mots monosyllabiques à intervalles réguliers, il produit en fait les mêmes types de déviations temporelles acoustiques que celles manifestées par les auditeurs lorsqu'ils cherchent à percevoir une séquence comme isochrone (ce dont parlait Morton et al., 1976). Fowler pose alors l'hypothèse que le corrélat du *p-center* se trouverait au niveau des gestes articulatoires : les déviations élicitées par rapport à une isochronie acoustique contrainte seraient systématiques et reflèteraient les différences de mode d'articulation des consonnes initiales de syllabe. Ainsi, une séquence de mots monosyllabiques identiques (par exemple : « *mad mad mad mad mad mad* ») que le locuteur a pour consigne de prononcer de manière régulière présentera des intervalles équivalents entre les attaques acoustiques des mots, puisque la consonne initiale ne varie pas ; tandis qu'une séquence de mots monosyllabiques

commençant par des consonnes différentes (par exemple : « *mad sad mad sad mad sad* ») présentera des intervalles acoustiquement irréguliers entre les attaques des mots. Du côté de l'auditeur, le fait que celui-ci considère comme isochrone cet énoncé présentant des variations par rapport à l'isochronie acoustique suggère qu'il localise les *p-centers* en se basant sur les indices acoustiques dans l'énoncé reflétant la temporalité des gestes articulatoires du locuteur (par exemple, un silence associé aux consonnes occlusives) (Fowler, 1979). L'isochronie (au sens de *régularité*) de la parole se trouverait alors plutôt au niveau de la programmation du rythme (Astésano, 2001, p. 47). En ce sens, la proposition de Fowler rejoint celle de la théorie motrice de perception de la parole de Liberman et Mattingly (1985), sur laquelle nous reviendrons dans la partie 2.1.1.1.

De plus, les *p-centers* seraient associés à la voyelle accentuée (selon sa saillance ou bien son poids métrique) et pourraient se situer au plus tôt au niveau de son initiation articulatoire ou un peu plus tard au niveau de son approximation la plus proche de sa forme cible dans l'agencement du tractus vocal, les *p-centers* étant de plus affectés par le mode d'articulation de la consonne prévocalique. En raison du phénomène de coarticulation, l'initiation articulatoire de la voyelle pourrait se produire pendant la production de la consonne prévocalique. La réalisation de la forme cible de la voyelle se manifesterait par une stabilisation de l'effet de la voyelle sur le signal acoustique de la consonne (Fowler, 1979). L'avantage du *p-center* comme unité accentuelle (par rapport notamment à la syllabe) réside dans la prise en compte de la nature des segments sur le plan co-articulatoire et de leur impact sur la perception accentuelle (Astésano, 2001, p. 65). Nous reviendrons sur les étapes de production de la parole et notamment le contrôle articulatoire avec le modèle neurolinguistique DIVA, qui permet d'expliquer les différences de latence engendrées par des différences de complexité articulatoire dans la partie 2.1.2 (p. 42).

En phonologie métrique, la notion d'accent peut également renvoyer à un niveau d'organisation et de traitement perceptif sous-jacent : l'accent métrique, entité relevant de la phonologie (équivalent de *stress* en anglais) (Astésano, 2017, p. 27), qui serait traité au moyen de processus *bottom-up* ou *top-down*. En effet, la proéminence d'une syllabe peut résulter de sa saillance acoustique (processus *bottom-up*) mais aussi de son poids métrique (processus *top-down*) (Astésano & Bertrand, 2016). La perception et le traitement de proéminences dans la parole dépend de notre intériorisation du battement (pulsation) sous-jacent distribuant ces proéminences. Ainsi, les proéminences (acoustiques) qui ne tombent pas sur un battement attendu (proéminence cognitive) pourraient ne pas être perçues telles quelles (Handel, 1989,

cité par Astésano, 2001, p. 23). Ce postulat rejoint également celui de l'hypothèse du rebond attentionnel (ABH), sur laquelle nous reviendrons en partie 2.2.3 (p. 55).

Le débit et les pauses participent aussi à l'organisation temporelle des proéminences et à l'impression de pulsation rythmique de la parole. Ils ne sont pas uniquement générateurs d'une pulsation simple mais également d'une structure métrique périodique. Sur le versant de la production, le débit de parole est généralement mesuré en considérant le nombre de syllabes par seconde. Il dépend de facteurs individuels ou situationnels et peut impacter la réalisation accentuelle en surface en instaurant de nouvelles contraintes physiologiques. Sur le plan rythmique, les pauses, remplies ou silencieuses, ont entre autres une fonction de démarcation entre unités rythmiques et interviennent donc lors de la segmentation du signal, en perception comme en production. De plus, la durée des pauses participe à la perception de régularité dans le discours en installant un battement silencieux (Astésano, 2001, p. 98, 100, 102).

Au-delà des schèmes métriques propres aux différentes langues, la succession d'accents dans la parole continue et instanciée de façon linéaire (signal de parole) doit respecter le principe universel d'eurythmie ou d'alternance optimisée. Pour éviter au maximum les collisions accentuelles (succession de deux battements forts) et les vides accentuels (succession trop longue de battements faibles), les accents sont amenés à se produire à intervalles quasi réguliers. Dans cette optique eurythmique universelle, la proéminence accentuelle revient ainsi toutes les trois à quatre syllabes, correspondant à un intervalle d'environ 600 ms, que l'on peut rapprocher de la période de référence (cf. 1.2.1.3.1, p. 19) (Di Cristo, 2003). Afin d'être perçues au bon endroit par l'auditeur, les proéminences produites par le locuteur doivent donc se plier à la fois aux contraintes eurythmiques de réalisation effective en surface, mais aussi aux contraintes métriques sous-jacentes liant proéminences à d'autres niveaux d'organisation linguistiques (unités lexicales, syntaxiques, sémantiques et informationnelles) (Astésano, 2017, p. 76).

1.2.2 L'APPROCHE DE LA PHONOLOGIE METRIQUE

Les mécanismes de segmentation en groupements et d'extraction de la pulsation ainsi que leur intégration dans des structures hiérarchiques sont au cœur de l'approche de la phonologie métrique : les règles phonologiques ne sont pas envisagées selon un enchaînement dans un ordre linéaire mais comme une organisation hiérarchique en niveaux de groupements (ou constituants) et niveaux de proéminences (Di Cristo, 2003).

Dans cette partie, nous commencerons par présenter les principes universaux de la prosodie (cf. 1.2.2.1) puis poursuivrons avec les représentations formelles de la structure prosodique (cf. 1.2.2.2, p. 29), avant de voir le cas plus spécifique du français (cf. 1.2.2.3, p. 30).

1.2.2.1 PRINCIPES GENERAUX : ACCENTUATION ET NIVEAUX DE CONSTITUANCE

Les accents, ou proéminences, correspondent aux éléments discrets, figures, qui émergent du fond sonore constitué par les syllabes inaccentuées, en faisant varier les paramètres structurels du signal sonore (durée, timbre, hauteur, intensité).

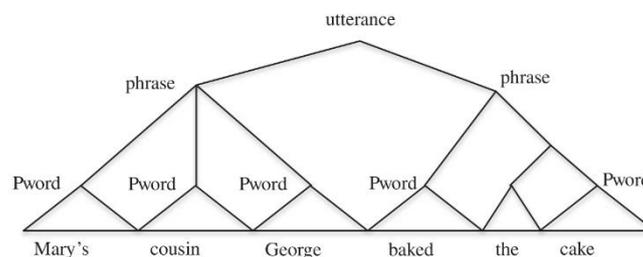
L'accentuation rythmique se distingue de l'accentuation emphatique (sémantico-pragmatique), cette dernière résultant d'un choix du locuteur d'attirer l'attention de l'auditeur sur un élément informatif ou affectif de l'énoncé et dont la place n'est pas contrainte en français. L'accent rythmique est multifonctionnel et intègre différents types de rôles : il est réputé assurer une fonction hiérarchique marquant la structure thématique ou informationnelle de l'énoncé, une fonction démarcative indiquant les frontières de mots ou syntagmes, une fonction contrastive par la mise en valeur des unités accentuables (syllabes), une fonction rythmique qui assure l'alternance optimale de syllabes accentuées et inaccentuées, une fonction culminative pour les langues dites accentuelles où chaque mot lexical porte un accent, et enfin, une fonction distinctive lorsque l'accentuation permet de discriminer entre des paires minimales métriques. Les accents rythmiques (initiaux et finaux) participent ainsi à la structuration rythmique de la parole en générant des groupements de sens, de manière à faciliter le décodage de la parole pour l'auditeur (Astésano, 2001). Seuls les mots lexicaux (substantifs, adjectifs, verbes) sont susceptibles de porter l'accentuation rythmique, contrairement aux articles, propositions et pronoms, structurellement non accentogènes (Di Cristo, 2016).

En tant que proéminence, l'accent distingue les battements (ou syllabes) selon qu'ils soient forts ou faibles et hiérarchise ces différences selon leur relative saillance en les associant à un niveau de constituance (groupement rythmique). Cette structure en constituants prosodiques s'élabore en interface avec la structuration syntaxique mais ne lui est pas isomorphe. Elle comprend plusieurs niveaux hiérarchiques où le niveau supérieur englobe le ou les constituants de niveau inférieur (Turk & Shattuck-Hufnagel, 2014).

D'autre part, l'effet des frontières de groupes (ou constituants) prosodiques sur la planification phonétique varie selon le niveau de constituance : un allongement associé à une

frontière de niveau de constituance supérieur sera plus important qu'un allongement associé à un niveau inférieur (Turk & Shattuck-Hufnagel, 2014). Par exemple, dans la phrase « *Mary's cousin George baked the cake* » (cf. **Figure 1**), l'allongement final associé au mot « *cousin* » (niveau de constituance du mot prosodique¹⁹ – *Pword* sur la figure) est moins important que celui associé au mot « *George* » (niveau de syntagme – *phrase*), lequel est moins important que l'allongement final associé au mot « *cake* », ce dernier mot étant associé à la frontière finale du constituant du niveau de l'énoncé (*utterance*), supérieur aux niveaux du mot prosodique et du syntagme.

Figure 1 :
Exemple de structure prosodique



Note. Exemple de structure prosodique avec les niveaux de constituants prosodiques suivants : *Pword* (mot prosodique), *phrase* (syntagme) et *utterance* (énoncé). Tiré de Turk et Shattuck-Hufnagel (2014, p. 3).

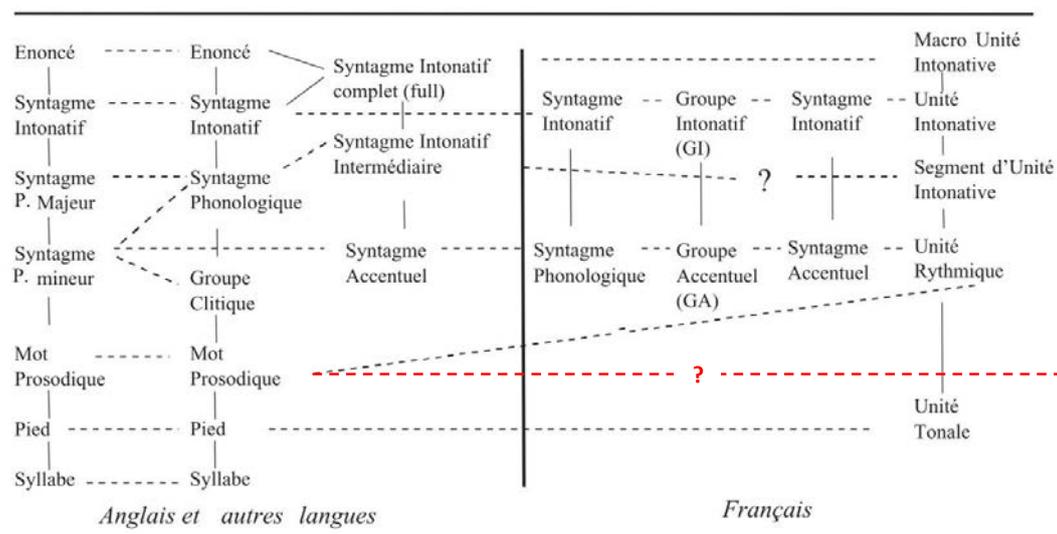
Les niveaux de constituance hiérarchiques font l'objet de débats selon les auteurs et selon les langues quant à leur nature, leur nombre et leur nom (Turk & Shattuck-Hufnagel, 2014) (cf. par exemple la **Figure 2**, tirée de Astésano, 2017). Par exemple en français, du fait qu'elle ne possède pas de fonction lexicalement distinctive, l'accentuation est traditionnellement considérée comme post-lexicale. C'est-à-dire qu'elle s'appliquerait non pas au niveau lexical (le mot) mais au-delà, au niveau du groupe de mots. De fait, deux niveaux de constituance prosodique sont généralement reconnus : le niveau du syntagme accentuel (niveau inférieur) et le niveau du syntagme intonatif (niveau supérieur). C'est au niveau du syntagme accentuel que se joueraient les règles phonologiques d'eurythmie et d'assignation des prééminences (accentuation initiale et finale), tandis que le syntagme intonatif est le lieu d'actualisation du contour intonatif et est constitué d'un ou plusieurs syntagmes accentuels (Astésano, 2017). Une proposition d'Astésano (2017) suggère cependant que, pour le français, le plus petit niveau d'actualisation de l'accentuation se situe à un niveau inférieur (groupement plus petit) au syntagme accentuel, au niveau du mot prosodique (ou mot phonologique). Celui-

¹⁹ Le mot prosodique renvoie à un groupement rythmique constitué d'un mot lexical éventuellement accompagné de son clitique (Selkirk, 1996, citée par Astésano, 2022).

ci correspond à un groupement constitué d'un mot lexical éventuellement accompagné de son déterminant (nous reviendrons sur l'existence du niveau du mot prosodique comme niveau d'actualisation de l'accentuation bipolaire en français dans la partie 1.2.2.3.1, p. 31). Certains auteurs intègrent également un niveau intermédiaire entre les syntagmes accentuel et intonatif, dont la pertinence pour le français est encore débattue (pour une discussion sur les niveaux de constituance en français, voir Astésano, 2017) (cf. **Figure 2**). Dans la ligne des travaux d'Astésano (2017), nous adoptons donc le modèle en niveaux de constituance suivants dont la pertinence a été démontrée de façon robuste pour le français : syllabe, pied, mot prosodique, syntagme accentuel, syntagme intonatif et énoncé. En fonction de la longueur (et donc de la profondeur hiérarchique) de l'énoncé, tous les niveaux peuvent ne pas être décrits de façon nettement différenciée dans la hiérarchie.

Figure 2 :

Terminologie utilisée dans les approches de la hiérarchie des domaines prosodiques



Note. Adapté de Astésano (2017, p. 38), reprenant l'illustration de Di Cristo (2011, p. 76). Le « ? » dans la partie supérieure renvoie au questionnement de Di Cristo (2016) à propos de l'absence du niveau du syntagme intermédiaire dans les modèles du français. Astésano interroge ici la présence du niveau du mot prosodique en français, qu'elle indique par un « ? » rouge.

Afin d'illustrer les différents niveaux de constituance et de proéminence, nous présentons dans la partie suivante les deux types de représentation formelle sous-jacente proposés par la phonologie métrique pour rendre compte des relations entre niveaux d'organisation hiérarchique accentuelle : la représentation en arbre métrique et la représentation en grille métrique.

1.2.2.2 REPRESENTATIONS DE LA STRUCTURE PROSODIQUE

La représentation en arbre métrique formalise la structure hiérarchique et relationnelle accentuelle par des nœuds et des branchements (cf. partie inférieure de la **Figure 3**). Chaque syllabe (s) est rattachée à un nœud de l'arbre et étiquetée « F » (Fort) pour une force métrique forte ou « f » (faible) pour une force métrique faible, cette propriété étant relationnelle : une syllabe est « F » par rapport à l'autre ou aux autres du même constituant et réciproquement pour une syllabe « f ». Plusieurs nœuds du même constituant sont ensuite reliés par des branchements à un nouveau nœud correspondant à une proéminence de niveau supérieur (par exemple, le niveau du pied) étiquetée f ou F, et ainsi de suite jusqu'au niveau de constituance de l'énoncé. Au premier niveau de constituance (syllabe), les nœuds « F » renvoient aux syllabes accentuées (accents initial et final – AI et AF²⁰), et les nœuds « f » aux syllabes inaccentuées. Aux niveaux supérieurs, les nœuds « F » et « f » se distinguent ensuite entre eux par leur force métrique relativement aux autres lexèmes du même constituant, et relativement à leur niveau dans la hiérarchie : un nœud « F » du niveau de l'énoncé a un niveau de proéminence supérieur à un nœud « F » du niveau du pied (Di Cristo, 2016).

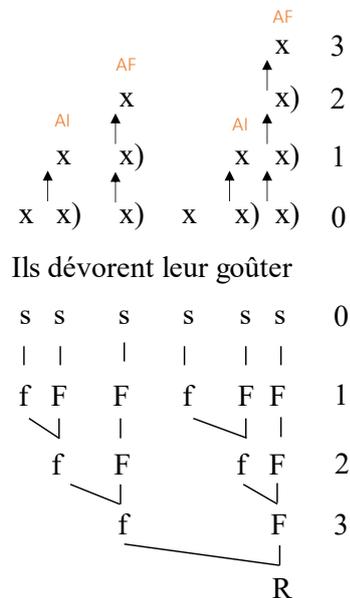
La représentation en grille métrique formalise le degré de force métrique d'une proéminence selon sa place assignée à un niveau inférieur ou supérieur dans la grille (cf. partie supérieure de la **Figure 3**). A chaque niveau de la grille, des groupements métriques (constituants) sont formés à partir des syllabes (« x »), chacun de ces groupements ne comportant qu'une seule tête métrique (syllabe la plus proéminente), étiquetée « x ». Les têtes métriques sont ensuite projetées en nouveaux groupements au niveau de constituance supérieur ne comportant à nouveau qu'une tête métrique (« x »), alignée sur la plus forte proéminence du niveau inférieur, et ce jusqu'au niveau de l'énoncé. Suivant le niveau auquel elle se situe, la proéminence est ainsi plus ou moins forte. Les « x » indiquent les frontières droites de constituant. Dans notre exemple **Figure 3** (« Ils dévorent leur goûter »), sur la ligne 2 (niveau de constituance du syntagme accentuel), la proéminence associée à la syllabe finale du mot prosodique « Ils dévorent » est moins forte que celle associée à la syllabe finale du mot

²⁰ AF est primaire par rapport à AI (secondaire), car AF présente une force métrique supérieure à AI.

prosodique « leur goûter », cette dernière se projetant alors au niveau supérieur du syntagme intonatif (ici indifférencié par rapport au niveau de l'énoncé).

Figure 3 :

Représentation de la structure métrique de l'énoncé « Ils dévorent leur goûter »



Note. Le syntagme « Ils dévorent leur goûter » est extrait du matériel expérimental de Cason, Astésano et al. (2015), que nous réutilisons dans le cadre de ce travail de thèse. En bas : arbre métrique (s : Syllabe ; f : battement faible ; F : battement fort ; R : Racine de l'arbre). En haut : grille métrique. Les « x » renvoient aux syllabes proéminentes. 0 : interface entre le plan segmental et le plan métrique (syllabes) ; 1 : pied ; 2 : syntagme accentuel ; 3 : syntagme intonatif. AI : Accent Initial ; AF : Accent Final. Adapté de Di Cristo (2016, p. 61).

Ainsi, par sa structure en branchements, l'arbre métrique met davantage en évidence les relations entre constituants métriques par les nœuds qui les relient, tandis que la grille métrique met davantage en évidence les propriétés hiérarchiques de l'accentuation, en indiquant que les proéminences des niveaux supérieurs sont perçues plus fortes que les proéminences des niveaux inférieurs. De plus, la grille métrique a l'avantage par rapport à la représentation en arbre d'intégrer les phénomènes d'eurythmie (suppression, insertion et déplacement de l'accent) (Astésano, 2001, p. 86, 87).

1.2.2.3 LA PROSODIE DU FRANÇAIS

La prosodie du français repose sur les principes de niveaux de proéminence et de constituance, qu'elle complète de principes plus spécifiques : le principe de bipolarisation accentuelle et le principe de dominance à droite (cf. 1.2.2.3.1). Nous avons vu dans la partie 1.2.1.3.2 que toutes les langues n'utilisent pas les mêmes variations structurelles pour distinguer les proéminences dans le flux de parole et présentons les caractéristiques phonétiques des accents rythmiques du français en partie 1.2.2.3.2.

1.2.2.3.1 PRINCIPES GENERAUX

En ce qui concerne le français, la métrique est caractérisée par deux principes fondamentaux : le principe de bipolarisation accentuelle, selon lequel la syllabe initiale et la syllabe finale d'un groupement peuvent recevoir un accent et le principe de dominance à droite, appliqué à tout niveau de constituance, qui place la tête métrique²¹ d'un groupe accentuel sur la syllabe à droite. Ainsi, le pied métrique²² le plus naturel est iambique en français (« xX ») : la tête métrique est la dernière syllabe²³ (Di Cristo, 2003). Ce paramètre de dominance à droite est cependant nuancé par des résultats d'Astésano (2017) dans sa proposition de pied spondéïque comme pied métrique élémentaire. Au cours des dernières décennies, les études comportementales sur l'accentuation du français tendent à montrer que l'accent initial (AI) et l'accent final (AF) seraient perçus de manière plus saillante que les syllabes inaccentuées, et ce quel que soit le niveau de constituance – et notamment dès le niveau du mot lexical (ou mot prosodique) (Astésano, 2001, 2017 ; Garnier, 2018). De plus, AI et AF seraient perçus au niveau lexical indépendamment de la force ou même de la présence de leurs corrélats acoustiques respectifs (pour l'AI, Jankowski et al., 1999 ; Te Rietmolen, 2019 ; pour l'AF, Méndez et Astésano, 2017 ; Garnier, 2018). Cela pourrait résulter d'un processus cognitif *top-down* qui contraindrait la perception des auditeurs. Les résultats d'études en EEG vont également dans ce sens (Aguilera et al., 2014 ; Astésano et al., 2013 ; Te Rietmolen et al., 2016). De plus, la présentation d'un patron métrique inattendu semble altérer l'accès au sens, en témoigne l'élicitation d'une composante N400 associée à un temps de latence et un taux d'erreur plus élevés (Magne et al., 2007), ce qui indique que la représentation du patron métrique est liée à la représentation phonologique du mot lexical, qui renvoie au niveau de constituance du mot prosodique.

L'ensemble de ces résultats conforte l'ancrage métrique de AI et AF dans la représentation phonologique (lexicale) des auditeurs francophones et invite à considérer le niveau du mot prosodique (ou mot phonologique) comme domaine d'actualisation de l'accentuation bipolaire du français, niveau inférieur au syntagme accentuel. Le pied métrique élémentaire du français ne serait alors plus iambique (xX) mais spondéïque (X-X), marqué à la

²¹ La tête métrique correspond à la syllabe accentuée dans la représentation sous-jacente.

²² Pied métrique : « unité formée d'une syllabe accentuée et d'un nombre optionnel (mais limité) de syllabes inaccentuées » (Di Cristo, 2003, p. 7-8).

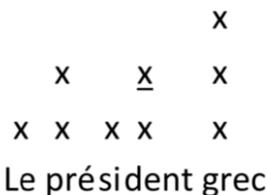
²³ A l'intérieur du mot, seul le schwa (e muet) ne peut recevoir d'accent, impliquant dans ce cas particulier un déplacement de l'accent final sur l'avant dernière syllabe si le schwa est en finale (Dell, 1984, p. 100).

fois par AI et AF, de force métrique comparable mais fonctionnellement différents (Astésano, 2017, 2019 ; Te Rietmolen, 2019).

En considérant d'autres travaux, nous retenons également que la liaison entre un accent initial rythmique et un accent final d'unité lexicale ou de syntagme forme une unité rythmique de niveau supérieur, l'arc accentuel. Cette unité assure une fonction intégrative d'unités rythmiques inférieures ainsi qu'une fonction démarcative de groupes rythmiques (Fonagy, 1979, cité par Di Cristo et Hirst, 1993). L'arc accentuel peut ainsi se situer à différents niveaux de la hiérarchie prosodique. Au-delà du niveau du mot prosodique, l'arc accentuel délimitant un groupe rythmique s'accompagne d'un autre phénomène, la désaccentuation : afin d'éviter les *clashes* (collisions) accentuels dans le groupe ou bien pour renforcer l'unité sémantique du groupe, l'accent final du ou des mots prosodiques internes au groupe s'efface partiellement au profit de l'arc accentuel (cf. **Figure 4**) (pour une discussion sur l'arc accentuel, se référer à Astésano et Bertrand, 2016).

Figure 4 :

Illustration de l'arc accentuel



Note. Grille métrique du syntagme « Le président grec ». La deuxième ligne en partant du haut illustre le phénomène de désaccentuation (x souligné) lié à un arc accentuel entre un accent initial (syllabe initiale de « président ») et un accent final (« grec »). Adapté de Astésano et Bertrand (2016, p. 16).

1.2.2.3.2 CARACTERISTIQUES PHONETICO-ACOUSTIQUES DES ACCENTS RYTHMIQUES

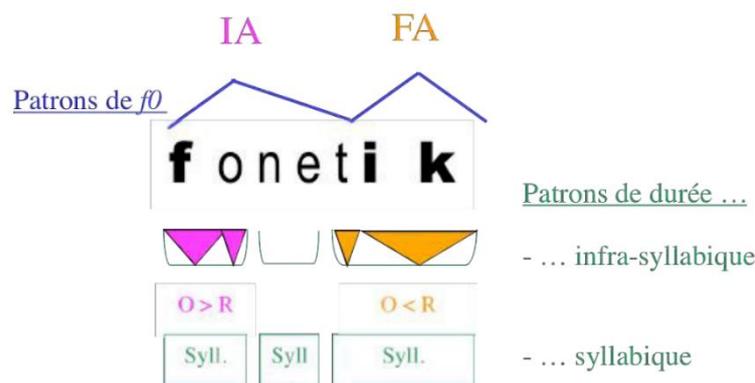
Tout comme nous l'avons vu dans la partie 1.2.1.3.2 (p. 21), les caractéristiques structurelles des accents rythmiques varient selon les langues.

Le français a longtemps été considéré comme une langue sans rythme accentuel, plutôt caractérisée par un rythme syllabique (Rossi, 1980, cité par Astésano, 2001). Cette faible perception de proéminence en français par rapport aux langues dites accentuelles pourrait être liée aux paramètres constitutifs de la structure rythmique sonore : la durée, la hauteur, l'intensité et le timbre. En français, l'accent final se manifeste principalement par une différenciation de durée ; en revanche, le timbre et l'intensité sont relativement stables en français, donc non distinctifs. Quant au paramètre de hauteur (ou fréquence fondamentale, « f0 »), il se manifeste sur un empan syllabique large, le rendant moins saillant perceptivement (Astésano, 2001, p.

253). L'accentuation rythmique en français est donc marquée par des différences touchant en particulier les paramètres de durée et de hauteur, représentatives de la distribution accentuelle.

L'accent initial en français se distingue par une configuration tonale asymétrique étendue sur plusieurs syllabes et un allongement infrasyllabique de l'attaque consonantique ; tandis que l'accent final présente une configuration tonale symétrique localisée sur la syllabe accentuée et un allongement infrasyllabique de la rime (Astésano, 2001, 2016) (cf. **Figure 5**). Il est important de noter que les durées segmentales (niveau infrasyllabique) sont soumises à différents types de variabilité structurelle, qu'il convient de prendre en compte dans l'analyse afin de faire émerger la variabilité systématique liée aux phénomènes rythmiques d'accentuation : variabilité due (1) aux phénomènes articulatoires, (2) au contexte phonétique et (3) au contexte prosodique (Astésano, 2001, p. 170).

Figure 5 :
Caractéristiques phonétiques de AI et AF en français



Note. En haut en bleu : configurations tonales de AI (IA) (pentes de f_0 asymétriques) et AF (FA) (pentes de f_0 symétriques). Les trois lignes du bas indiquent les patrons de durée au niveau syllabique : AF > AI > syllabe inaccentuée et au niveau infrasyllabique : attaque (O - onset) > rime (R) pour AI et rime > attaque pour AF. Tiré de Astésano (2017, p. 83).

Selon Astésano (2001), les caractéristiques phonétiques spécifiques de l'accent initial (AI) et de l'accent final (AF) du français (concernant les paramètres de durée ou de hauteur) amènent à les considérer tous deux comme accents phonologiques constitutifs de la structure sous-jacente du mot, et bornes de l'arc accentuel. Ces formes sous-jacentes constituent les « points d'ancrage » d'une langue donnée, dont les « alloformes » (variations) se présentent comme des révélateurs de facteurs contextuels ou contraintes pragmatiques, notamment des styles de parole. Par exemple, au niveau infrasyllabique des prééminences, la parole lue aura tendance à présenter un allongement des codas, tandis que la parole spontanée aura tendance à marquer davantage l'allongement des attaques.

Au niveau syllabique (et non plus infrasyllabique), l'allongement de l'accent final en français par rapport à l'accent initial va intensifier subjectivement la sensibilité de sa perception, le rendant plus saillant que l'accent initial. De plus, l'accent initial est davantage soumis à des contraintes phonotactiques et d'eurythmie. Pour ces raisons, l'accent initial a pu être qualifié de secondaire face à l'accent final, plus systématique et donc primaire (Di Cristo & Hirst, 1993). Dans ce travail de thèse, nous nous sommes plutôt concentrée sur la manipulation de l'accent primaire (AF). En effet, notre expérience d'amorçage rythmique de la parole manipule la congruence métrique (= la similarité de structure métrique sous-jacente) entre une amorce rythmique et une cible langagière, en faisant varier l'emplacement de l'AF (cf. partie 3.2, p. 89).

1.3 SYNTHÈSE : UNE PAROLE RYTHMIQUE

Dans ce premier chapitre, nous avons cherché à comprendre et définir la rythmicité de la parole à partir de différentes approches permettant d'en appréhender les structures sur les plans phénoménologique, phonologique métrique, acoustico-phonétique et prosodique. Dans un premier temps, nous avons situé le concept général de 'rythme' du point de vue de la philosophie des phénomènes rythmiques : selon Sauvanet (2000), le rythme n'existe pas par lui-même en tant qu'objet mais en tant que phénomène rythmique perçu, subi ou agi par un sujet. Un objet est ainsi rythmique dès lors qu'il est caractérisé par la combinaison d'au moins deux critères parmi les trois suivants : une Structure (contrastive), une Périodicité (successive, prévisible) et un Mouvement (vivant, imprévisible).

Dans un second temps, nous avons appliqué ces critères à la parole rythmique que nous avons ensuite abordée suivant l'approche de la phonologie métrique (Astésano, 2001, 2017 ; Di Cristo, 2003), laquelle est construite sur les mécanismes de traitement des séquences auditives complexes (Drake, 1998, 2010). Nous avons ainsi vu que le traitement de la parole est lié à deux processus de base nécessaires au traitement de l'information auditive et à l'intégration de ces processus dans des structures hiérarchiques : la segmentation du flux sonore en groupements (niveaux de constituance hiérarchiques) et l'extraction d'une pulsation (niveaux de prééminence hiérarchiques). De manière intéressante, ces deux processus prennent appui sur les trois critères de rythmicité, le groupement se faisant selon des caractéristiques structurelles et l'extraction de la pulsation se faisant à partir de la succession de battements de pulsation ; le dernier critère de mouvement se retrouvant dans toute activité humaine. Nous avons vu que la notion d'isochronie dans la parole ne fait pas consensus et qu'elle n'est pas nécessaire pour que soit générée une sensation de pulsation rythmique. Dans le cadre de l'approche métrique, la sensation de pulsation est rattachée aux prééminences, les syllabes accentuées, qui s'organisent temporellement pour structurer la prosodie des énoncés. Ces prééminences peuvent être perçues par leur prégnance (saillance acoustique significative – processus *bottom-up*) ou par leur poids métrique (processus *top-down*).

Nous retenons ainsi que la parole, en tant que phénomène rythmique, peut être appréhendée selon diverses approches qui nous permettent de mieux cadrer notre objet d'étude tout en l'ancrant dans une perspective linguistique. Cependant, afin de poursuivre notre objectif de recherche en explorant l'impact de différents amorçages rythmiques sur la production de parole d'un point de vue comportemental, il nous faut à présent poursuivre notre cheminement

théorique en intégrant les modèles neuropsycholinguistiques de traitement de la parole et aborder les modèles dédiés aux mécanismes de synchronisation rythmique à une stimulation rythmique extérieure, c'est-à-dire à une amorce rythmique. Cela nous permettra ainsi de proposer dans le chapitre suivant un point de vue argumenté sur le rôle du rythme dans le traitement de la parole, en prenant en considération les aspects cognitifs en perception et production d'une part, et de synchronisation rythmique d'autre part.

2 LE RÔLE DU RYTHME DANS LE TRAITEMENT DE LA PAROLE

Maintenant que nous avons établi le caractère rythmique de la parole dans la partie précédente, nous pouvons nous demander de quelle manière un stimulus rythmique externe (parole ou musique) pourrait venir interagir avec le traitement de la parole. Pour cela, nous présentons dans ce deuxième chapitre les mécanismes cognitifs qui sous-tendent le traitement de la perception et de la production de la parole (cf. 2.1) ainsi que les mécanismes de synchronisation et *entraînement*²⁴ rythmiques qui participent de ce traitement (cf. 2.2, p. 49). Pour finir, nous verrons dans quelle mesure le caractère supramodal du rythme intervient dans le traitement du langage (cf. 2.3, p. 61).

2.1 MÉCANISMES DE TRAITEMENT DE LA PAROLE

L'amorçage rythmique de la parole implique une étape de traitement d'une amorce auditive suivie d'une étape de production d'une cible langagière. Dans cette partie, nous présentons ainsi les mécanismes de traitement de la parole, c'est-à-dire les mécanismes cognitifs qui nous permettent de percevoir et décoder un signal auditif de parole (cf. 2.1.1) mais également les étapes impliquées dans la programmation et l'exécution de la production de parole (cf. 2.1.2, p. 42).

2.1.1 PERCEPTION DE LA PAROLE

2.1.1.1 DIFFÉRENTES APPROCHES DE LA PERCEPTION DE LA PAROLE

En perception de la parole, les approches auditives, motrices et sensorimotrices coexistent.

Les approches auditives proposent que la perception de parole et donc la mise en correspondance d'un son et d'un phonème se fasse par le biais des invariants acoustiques du signal sonore. Cette théorie est notamment appuyée par l'effet d'équivalence motrice : un même phonème (/u/ par exemple) peut être prononcé *via* différentes configurations articulatoires (Grabski & Sato, 2016). De plus, le fait que les nouveau-nés montrent des capacités précoces

²⁴ L'*entraînement* fait référence à la synchronisation graduelle entre un oscillateur et un rythme externe : la période initiale de l'oscillateur est entraînée par celle du rythme externe (Port et al., 1995).

de perception catégorielle auditive sans déjà maîtriser la production de parole laisse également penser que ceux-ci se reposent sur les indices acoustiques pertinents (Vaissière, 2020).

Opposée aux théories auditives, la théorie motrice de Liberman et Mattingly (1985) suggère que l'auditeur ne s'appuie pas sur les indices acoustiques seuls pour la perception phonétique. Deux postulats sont émis : en premier lieu, la perception phonétique serait une perception des gestes phonétiques ; en second lieu, la perception et la production seraient si intimement liées qu'elles correspondraient plutôt à deux faces d'une même unité phonétique. En effet, ces auteurs affirment que les indices acoustiques ne sont pas suffisants pour rendre compte de la catégorisation phonétique opérée par l'auditeur : en raison de la coarticulation, un même indice acoustique peut être perçu (donc catégorisé) différemment dans un contexte différent (ou inversement), ou encore, l'absence d'un indice acoustique pertinent peut être compensée par la présence d'autres indices (par exemple, le dévoisement d'une voyelle chuchotée). Les objets de la perception phonétique de la parole seraient non plus seulement de nature acoustico-phonétique, mais renverraient aussi pour une grande part aux gestes intentionnels du locuteur, lesquels correspondent aux commandes motrices invariantes des mouvements articulatoires, associés à des configurations linguistiquement significatives. Le signal acoustique servirait dans ce cas de source d'information à propos des gestes. La perception phonétique se différencierait ainsi de la perception auditive purement acoustique. Ces deux modes de perception pourraient être associés à deux modules²⁵ spécifiques distincts en compétition : à partir d'une même entrée acoustique, deux objets de perception seraient produits, l'un auditif (représentant les propriétés physiques acoustiques du son) et l'autre phonétique (représentant les gestes intentionnels du locuteur). Si les auditeurs et locuteurs traitent la parole par le même ensemble d'invariants moteurs (les gestes phonétiques), perception et production de la parole seraient intimement liées, permettant une conversion directe automatique du signal acoustique en gestes phonétiques (Liberman & Mattingly, 1985). Cette approche est notamment soutenue par des études de potentiels évoqués montrant l'activation de régions corticales motrices de production de la parole lors de la perception de parole (Fadiga et al., 2002 ; Watkins et al., 2003).

Cependant, la théorie motrice ne permet pas de rendre compte des effets d'équivalence motrice exposés précédemment, selon lesquels différentes configurations gestuelles renvoient au même son : il semblerait plutôt que ce soit la perception auditive qui influencerait les

²⁵ Un 'module' s'applique à un domaine spécifique et est associé à une architecture neuronale spécialisée (Liberman & Mattingly, 1985).

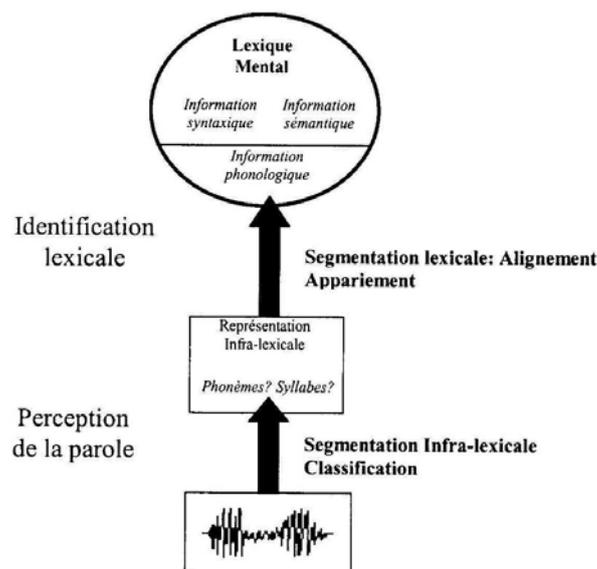
représentations abstraites motrices plutôt que l'inverse. La théorie de la perception pour le contrôle de l'action (PACT²⁶) (Schwartz et al., 2012) propose une alternative sensorimotrice et postule une co-structuration développementale entre perception sensorielle et production motrice de la parole (Vilain et al., 2019). De plus, la parole ayant une visée communicative, la perception de la parole serait contrainte par les limites motrices de production de la parole, tout comme la production de la parole qui, elle aussi, serait contrainte par les limites sensorielles de perception de la parole.

2.1.1.2 MODELES NEUROPSYCHOLINGUISTIQUES DE PERCEPTION DE LA PAROLE

Bagou et Frauenfelder (2002) proposent un modèle simplifié du traitement lexical afin d'associer une entrée sensorielle à des représentations lexicales mentales. Dans leur modèle, la perception de la parole consiste en différentes opérations et niveaux de représentations, parmi lesquelles les opérations de segmentation du flux sonore et de classification, les opérations d'identification lexicale et les représentations infra-lexicale et lexicale (cf. **Figure 6**). Afin de segmenter la parole continue et donc d'identifier correctement le début et la fin des unités linguistiques, l'interlocuteur se repose sur trois indices principaux : les indices phonétiques (acoustiques), les indices phonologiques (segmentaux et supra-segmentaux) et l'information lexicale.

Figure 6 :

Modèle simplifié de la reconnaissance des mots de Bagou et Frauenfelder (2002)



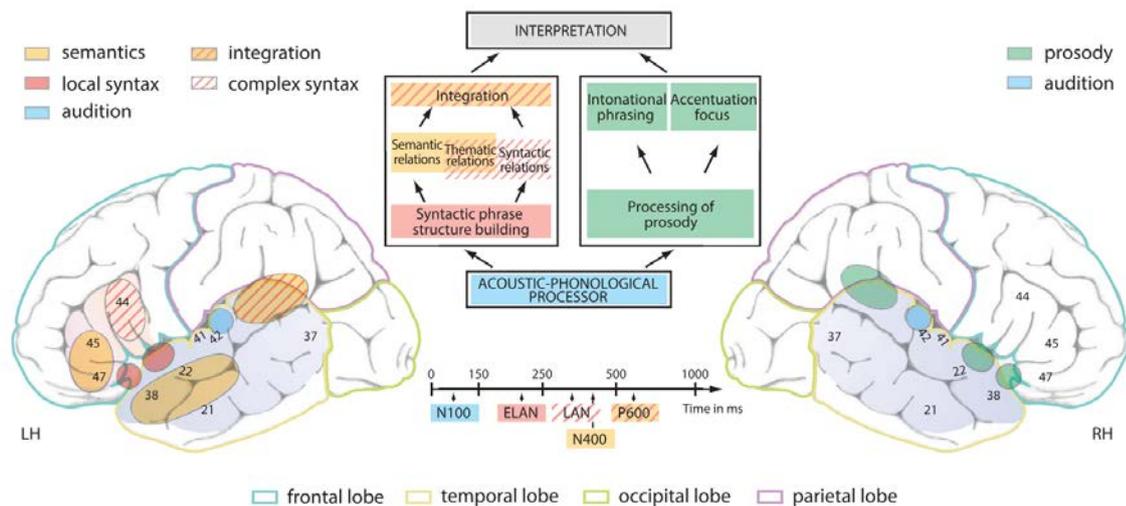
Note. Du signal d'entrée aux représentations lexicales. Tiré de Bagou et Frauenfelder (2002, p. 69).

²⁶ *Perception for Action Control Theory.*

Le modèle de Friederici (2017) intègre les données neurolinguistiques publiées dans une vaste littérature en perception/compréhension du langage pour préciser ces étapes de traitement et les associer à leur signature spatiale et temporelle de traitement, notamment sur la base d'études de potentiels évoqués (cf. **Figure 7**). Le traitement acoustico-phonologique prend place aux alentours de 100 ms, suivi du traitement des processus initiaux de construction syntaxique (~120 à 250 ms), du traitement morpho-syntaxique (~300 à 500 ms), du traitement lexico-sémantique (~ 400 ms) et du traitement des processus d'intégration syntaxique (~ 600 ms). Bien que ce modèle soit informatif du point de vue de la séquentialité du traitement de la parole, il demeure limité quant à l'intégration des processus prosodiques et notamment de l'accentuation, et ne considère pas l'implication du système moteur dans la perception de la parole. Cependant, des travaux récents, à l'instar de Te Rietmolen (2019), viennent apporter des éléments complémentaires pour permettre une meilleure compréhension du traitement temporel des proéminences. En effet, d'après Te Rietmolen (2019), l'établissement du patron métrique actualisé par l'accent initial (AI) et l'accent final (AF) se déroule aux alentours de 325 ms.

Figure 7 :

Modèle neurocognitif de la compréhension du langage de Friederici (2017)

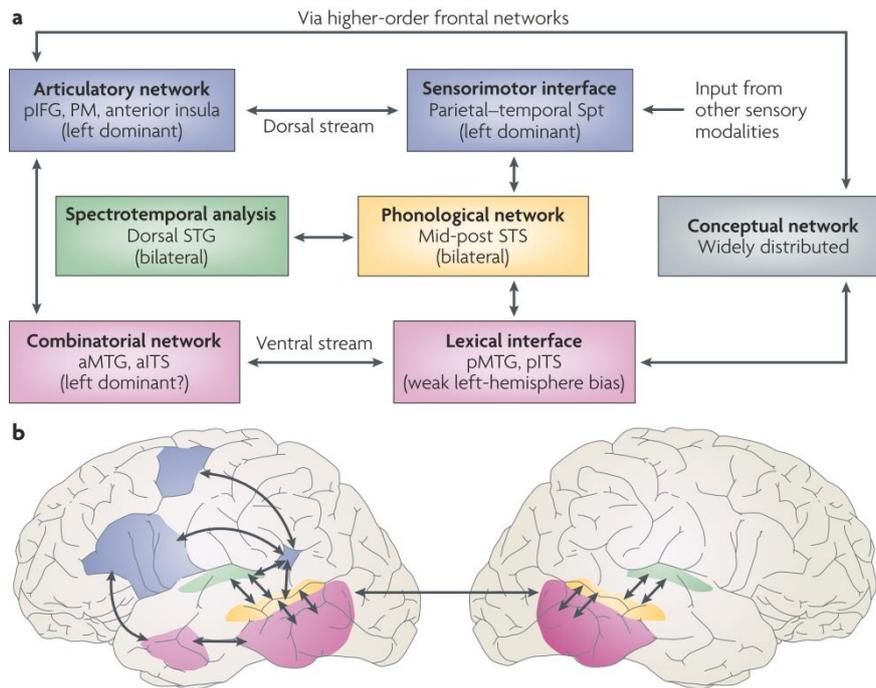


Note. Représentation de la neuroanatomie fonctionnelle du langage et ses phases de traitement. LH : hémisphère gauche ; RH : hémisphère droit. Tiré de Friederici (2017, p. 83).

La perception de parole s'accompagne de l'activation des régions motrices et proprioceptives de production de la parole (cf. l'approche PACT présentée dans la partie précédente 2.1.1.1, p. 37), certains modèles incluent un rôle des représentations motrices dans le décodage du signal de parole, en supposant un couplage fort sensori-moteur.

On retrouve ce type d'approche dans le modèle à double-voie (ventrale / dorsale) développé par Hickok et Poeppel (2004, 2007) qui proposent ainsi que le traitement du signal de parole impliquerait à la fois des représentations auditives et motrices (cf. **Figure 8**). Afin de justifier l'intégration d'une composante sensori-motrice dans leur modèle en argumentant en faveur d'un fort couplage entre sensorialité et motricité, les auteurs s'appuient sur (1) le développement du langage chez l'enfant, qui repose sur la production de cibles articulatoires de plus en plus précises au fur et à mesure de leur interaction avec l'input sensoriel de parole ; (2) les études lésionnelles montrant un impact sur la production d'une lésion touchant les régions auditives ; (3) les études d'imagerie fonctionnelle, qui mettent en évidence l'activation d'un réseau neuronal sensori-moteur dans la perception et la reproduction silencieuse de parole. Dans le modèle de Hickok et Poeppel, l'entrée sensorielle est d'abord traitée dans le cortex auditif puis le traitement se poursuit sur deux voies : une voie ventrale et une voie dorsale. La voie ventrale associe le son au sens (niveau de représentation du lemme) en se projetant ventro-latéralement jusqu'au cortex temporal antérieur (avec de possibles interactions avec les régions frontales) et la voie dorsale relie le son aux représentations motrices articulatoires en se projetant dorso-postérieurement jusqu'aux régions frontales. La voie ventrale emploie également deux routes parallèles afin de traiter conjointement les deux échelles temporelles convoquées pour la perception de la parole (cf. les régimes spectral et temporel évoqués par A. Albert et Grice (sous presse), présentés dans la partie 1.2.1.1, p. 15) : une première route utilise un réseau d'échantillonnage sur la bande de fréquences cérébrales gamma pour traiter les informations segmentales (au niveau du phonème) (~20-50 ms) et une seconde route utilise un réseau d'échantillonnage sur la bande thêta pour traiter les informations suprasegmentales (au niveau de la syllabe) (~150-300 ms).

Figure 8 :
Modèle à double voie ventrale/dorsale de Hickok et Poeppel (2007)



Note. (a) Schéma du modèle. (b) Localisation anatomique des composants du modèle. En bleu : la voie dorsale, en rose : la voie ventrale. a/pITS : sillon temporal inférieur antérieur/postérieur ; a/pMTG : gyrus temporal moyen antérieur/postérieur ; pIFG : gyrus frontal inférieur postérieur ; PM : cortex prémoteur ; STG : gyrus temporal supérieur ; STS : sillon temporal supérieur. Tiré de Hickok et Poeppel (2007, p. 395).

Les trois modèles présentés ici se complètent et nous renseignent sur les étapes de traitement d'un stimulus auditif de parole. Le modèle de Bagou et Frauenfelder (2002) nous présente le traitement de la parole comme résultant de deux étapes principales : la perception de la parole et l'identification lexicale. Le modèle de Friederici (2017) détaille ces étapes en les associant à leurs signatures temporelles : l'ensemble du traitement s'effectue ainsi en environ 600 ms. Finalement, le modèle de Hickok et Poeppel (2004, 2007) renforce l'intégration de l'activation des représentations motrices et multisensorielles et leurs circuits neuronaux sous-jacents. La perception de la parole s'accompagnerait ainsi de l'activation des aires motrices de production de la parole.

2.1.2 PRODUCTION DE LA PAROLE

La production de la parole est guidée par l'intention communicative du locuteur et donc par la génération de cibles de parole adaptées à la situation et en particulier à l'interlocuteur. La

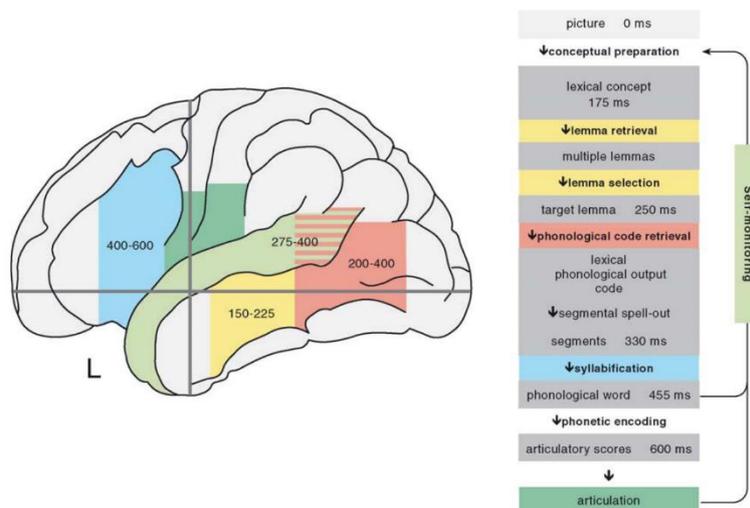
programmation motrice de la parole s'accompagne ainsi de l'émission de copies d'efférence²⁷ simulant la séquence sensorielle cible de parole à produire en fonction des gestes articulatoires envisagés. Un mécanisme de boucles régulatrices permet la comparaison en ligne (*online*) entre la séquence de parole produite et la copie d'efférence, afin d'éventuellement corriger ou adapter les commandes motrices pour atteindre la séquence cible (Grabski & Sato, 2016).

2.1.2.1 MODELES NEUROPSYCHOLINGUISTIQUES DE PRODUCTION DE LA PAROLE

Selon le modèle d'Indefrey et Levelt (2004), les étapes sérielles de production de la parole sont tributaires d'un temps de latence ou d'une initiation de la parole d'environ 600 ms à partir du moment où le stimulus est présenté (cf. **Figure 9**). La préparation conceptuelle s'effectue en ~175 ms, suivie de la récupération et de la sélection du lemme (+ ~75 ms = 250 ms), suivie de l'encodage phonologique (+ ~80 ms = 330 ms), suivie de la syllabification (+ ~125 ms = 455 ms) et de l'encodage phonétique (+ ~145 ms = 600 ms). Selon ce modèle, l'unité centrale de l'encodage phonétique est la syllabe, chaque syllabe de la langue d'un locuteur étant associée à des informations gestuelles (le syllabaire) et stockées dans le lexique mental (Levelt et al., 1999). Il suffit ensuite d'assembler les syllabes de manière linéaire pour produire un énoncé.

Figure 9 :

Modèle d'encodage lexical pour la production de parole d'Indefrey et Levelt (2004)



Note. Représentation schématique des aires cérébrales (à gauche) impliquées dans les étapes séquentielles (à droite) du modèle d'encodage lexical pour la production de la parole (dénomination d'image ou lecture de mot). En bleu : gyrus frontal inférieur postérieur (aire de Broca) ; en vert foncé : aires sensorimotrices ventrales ; en vert clair : gyrus temporal supérieur ; en jaune : gyrus temporal médian antérieur ; en rouge : gyrus temporal médian postérieur. Tiré d'Indefrey et Levelt (2004, p. 126).

²⁷ Une copie d'efférence est un « duplicata des commandes motrices » (Løevenbruck, 2019).

Contrastant avec la vision d'Indefrey et Levelt, Aichert et al. (2021) suggèrent que la planification phonétique ne s'effectuerait pas de manière sérielle à partir du syllabaire mais de manière hiérarchique s'étendant des gestes articulatoires aux structures métriques (portées par les syllabes). Selon ce modèle gestuel non-linéaire (*NonLinear Gestural – NLG model*), la réalisation phonétique est affectée par la position et le poids métrique des unités dans le patron métrique sous-jacent du mot, puis par les contraintes rythmiques de surface de réalisation de l'énoncé cible à produire (Aichert et al., 2021). D'après Aichert et al. (2021), deux stades sont en effet pertinents dans la production rythmique de parole : l'encodage phonologique (information métrique et segmentale) et l'encodage phonétique (récupération de la programmation motrice des gestes articulatoires). La portée du modèle reste cependant limitée, celui-ci ayant été testé sur peu de participants (participants avec apraxie) et seulement pour l'allemand (Ziegler et al., 2021).

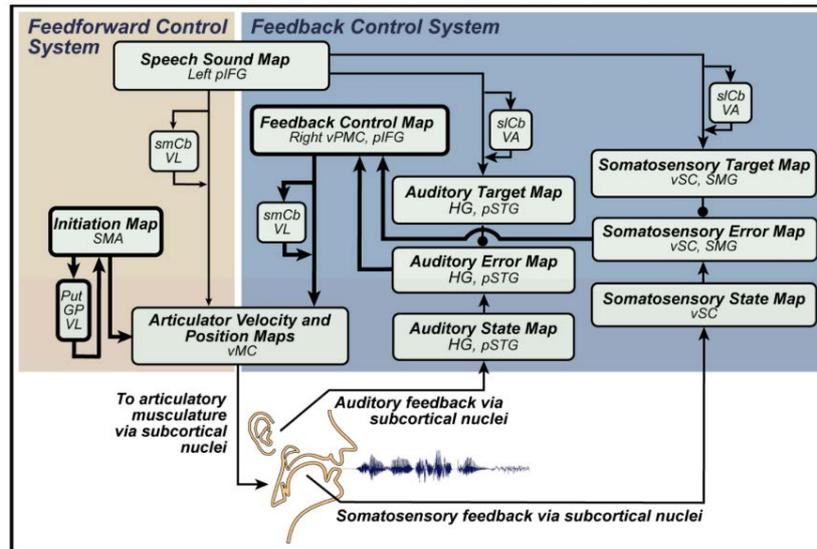
Les modèles de production de la parole peuvent encore être distingués selon qu'ils font appel ou non à un chronomètre extrinsèque (Turk & Shattuck-Hufnagel, 2014). En effet, la temporalité fait partie intégrante de la production de la parole, tant dans l'activité motrice coordonnée des organes de la phonation que dans la séquence sonore de sons produits.

Dans les modèles n'impliquant pas de chronomètre extrinsèque, les schémas temporels de surface émergent naturellement à partir des dynamiques internes du système de contrôle moteur de la parole. Selon Turk et Shattuck-Hufnagel, les modèles de contrôle moteur de la parole n'impliquant pas un chronomètre extrinsèque (ou reposant uniquement sur des oscillations intrinsèques) ne permettent pas de rendre compte pleinement du contrôle de la temporalité dans la parole. Elles avancent notamment comme argument en faveur d'un contrôle par un chronomètre extrinsèque le fait que la variabilité temporelle augmente avec la longueur des énoncés ou que des différences de variabilité sont associées à diverses parties d'un mouvement se déroulant dans le temps. Ces différences de variabilité temporelle selon la partie du mouvement observée impliquent un contrôle en temps réel dépendant de la tâche, des différences de priorité dans le traitement des parties observées et donc des traitements temporels indépendants : au plus près de la cible du mouvement, le traitement temporel est plus fin (moins variable), tandis que au plus loin de la cible du mouvement, le traitement est moins précis (plus variable). Par exemple, la théorie AP/TD (*Articulatory Phonology/Task Dynamics*) est basée sur des oscillateurs internes (unités internes au système) qui vont, par leurs caractéristiques, dicter la temporalité de surface (par exemple, la vitesse de parole ou la coordination des mouvements articulatoires).

En revanche, concernant les modèles impliquant un chronomètre extrinsèque, la temporalité résulte d'un contrôle explicite extérieur au système, opérant durant la planification de l'énoncé puis au cours de la production de l'énoncé. Parmi les modèles impliquant un chronomètre extrinsèque, on retrouve notamment le modèle DIVA (*Direction Into Velocities of Articulators*) (cf. **Figure 10**) et des modèles de la théorie du contrôle optimal. Ces modèles supposent qu'un mouvement généré répond à la loi du moindre effort : il doit avoir une trajectoire la plus optimale possible pour atteindre son objectif avec un coût minimal, tout en s'adaptant en temps réel à la tâche (Turk & Shattuck-Hufnagel, 2014).

Le modèle DIVA proposé par Tourville et Guenther (2011) est un modèle de production de la parole qui décrit les interactions sensorimotrices à l'œuvre lors du contrôle moteur articulaire. Ce modèle, toujours en développement, découle de l'avancée des travaux de recherche en imagerie cérébrale indiquant que les régions cérébrales impliquées dans la production de parole sont également celles impliquées dans la programmation et l'exécution de mouvements, et celles impliquées dans le traitement phonologique et acoustique des sons de la parole. Le modèle permet de simuler le contrôle du tractus vocal au moyen de deux systèmes de contrôle : un système *feedforward* (anticipation) et un système *feedback* (rétroaction). Chaque système de contrôle comprend plusieurs cartes neuronales associées à une fonction et un ensemble de neurones, qui interagissent afin de produire la parole. Par exemple, la carte d'initiation (*initiation map*) est associée à l'aire motrice supplémentaire (SMA) et contrôle les commandes motrices pour l'initiation de la parole.

Figure 10 :
Modèle DIVA d'articulation de la parole de Tourville et Guenther (2011)



Note. Modèle DIVA d'articulation de la parole. Abréviations: GP = globus pallidus; HG = gyrus de Heschl ; pIFG = gyrus frontal inférieur postérieur ; pSTG = gyrus temporal supérieur postérieur ; Put = putamen; sICB = cervelet latéral supérieur ; smCb = cervelet médial supérieur ; SMA = aire motrice supplémentaire ; SMG = gyrus supramarginal ; VA = noyau ventral antérieur du cervelet ; VL = noyau ventral latéral du thalamus ; vMC = cortex moteur ventral ; vPMC = cortex prémoteur ventral ; vSC = cortex somatosensoriel ventral. Tiré de Tourville et Guenther (2011, p. 23).

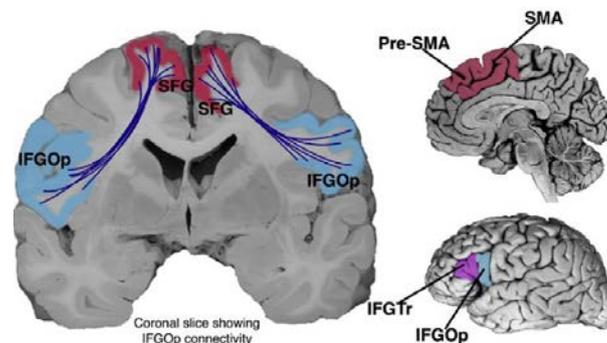
De plus, la SMA est reliée à l'aire de Broca²⁸ par le faisceau frontal aslant (FFA)²⁹ (cf. **Figure 11**), ce dernier étant impliqué dans la planification, la temporalité et la coordination motrice séquentielle bilatérale (Dick et al., 2019 ; Gatto et al., 2024) ainsi que les mouvements volontaires des mains (Budisavljevic et al., 2017). Plus spécifiquement, le FFA gauche jouerait un rôle dans l'initiation et la fluence de la parole et les mouvements manuels, tandis que le FFA droit jouerait un rôle dans les mécanismes généraux de contrôle de l'action et les fonctions exécutives (Dick et al., 2019). Une atteinte de ce faisceau se traduit par un trouble de la fluence verbale et se retrouve chez des personnes présentant une aphasia progressive de la parole (dégénérescence du FFA), une aphasia agrammatique ou un bégaiement (Gatto et al., 2024 ; Kemerdere et al., 2016 ; Kronfeld-Duenias et al., 2016).

²⁸ L'aire de Broca se situe dans le Gyrus Frontal Inférieur (IFI).

²⁹ Également appelé « Faisceau frontal oblique ».

Figure 11 :

Illustration des connexions formant le faisceau frontal aslant



Note. Illustration des connexions entre le gyrus frontal inférieur (IFGOp et IFGTr) et les aires motrices (pré-)supplémentaires ((Pre-)SMA) dans le gyrus frontal supérieur (SFG), qui forment le Faisceau Frontal Aslant. Tiré de Dick et al. (2019, p. 149).

L'implication du FFA dans le contrôle moteur des gestes manuels rejoint également l'hypothèse d'une origine motrice manuelle du langage oral, d'un point de vue phylogénétique et ontogénétique. Dans cette optique, le langage oral serait en premier lieu un système gestuel plutôt qu'acoustique (Gentilucci & Corballis, 2006), ce qui rejoint notamment les propositions de la théorie motrice de perception de la parole (cf. 2.1.1.1, p. 37). En effet, d'un point de vue phylogénétique, différentes théories s'opposent quant à l'origine du langage : origine vocale, origine gestuelle manuelle ou co-évolution des systèmes manuel et vocal. D'un point de vue ontogénétique, les étapes de production de gestes manuels sont corrélées à celles de la production vocale (Dohen et al., 2011). De plus, la réalisation de gestes manuels accompagnant la parole ne semble pas répondre (seulement) à un objectif communicatif pour faciliter le décodage, puisque ces gestes sont également effectués spontanément par des personnes aveugles communiquant entre elles (Iverson & Goldin-Meadow, 1998). Le partage anatomo-fonctionnel entre les gestes manuels et la production de la parole est notamment présent lors de l'activité de pointage, qu'elle soit manuelle, oculaire ou vocale prosodique (emphasis acoustique et articulatoire) : celle-ci convoque un réseau d'aires cérébrales commun, à savoir le lobe pariétal supérieur gauche. Plus précisément, ces trois types de pointage présentent en commun la nécessité d'une maîtrise fine des gestes (manuels, oculaires, articulatoires) qui transparait dans l'activation d'aires associatives dans le lobule pariétal supérieur, renforçant par là même le caractère multisensoriel du pointage et l'implication de représentations sensori-motrices (Dohen et al., 2011 ; Løevenbruck et al., 2009). La production gestuelle manuelle faciliterait

ainsi le décodage et l'encodage de la parole, en stimulant l'activité motrice dans les aires corticales de production de la parole (Iverson & Thelen, 1999).

Les deux modèles présentés sont complémentaires : le modèle d'Indefrey et Levelt (2004) se concentre principalement sur les étapes d'accès lexical jusqu'à la programmation motrice, tandis que le modèle DIVA (Tourville & Guenther, 2011) prend le relais et se concentre sur les processus articulatoires et sensorimoteurs. Ces deux modèles nous renseignent ainsi sur le déroulement temporel de la production de la parole, résultant en un temps d'initiation moyen de 600 ms, ainsi que sur le caractère sensorimoteur de la production *via* la mise en œuvre de boucles de contrôle (anticipation et rétroaction). Nous nous référons à cette variable de temps d'initiation de la parole dans notre activité expérimentale d'amorçage rythmique de la parole (cf. partie 3.2, p. 89), renseignant sur le caractère plus ou moins facilitant/perturbant d'une condition expérimentale par rapport à une autre. Pour finir, les travaux en neurolinguistique postulent également l'existence d'un partage anatomo-fonctionnel entre les régions motrices contrôlant les gestes manuels et les gestes de production de la parole. L'engagement moteur manuel pourrait ainsi avoir un impact sur les mécanismes d'initiation de la parole, ce qui constitue un des questionnements posés dans le cadre de notre étude.

2.2 *ÊTRE SUR LA MÊME LONGUEUR D'ONDE, UN PRÉREQUIS POUR LA COMMUNICATION*

Comme nous l'avons vu dans la partie 1.2.1 (p. 15), le signal de parole est une séquence auditive rythmique complexe. Il convoque donc, en plus des mécanismes cognitifs spécifiques au traitement du langage (cf. 2.1, p. 37), des mécanismes cognitifs plus généraux du traitement attentionnel. Ceux-ci vont nous permettre d'expliquer comment l'auditeur ou le locuteur réussissent à traiter aussi efficacement que possible le flux de parole linéarisé dans le temps. En effet, afin de décoder le signal de parole qui lui parvient, l'auditeur doit pouvoir le segmenter en groupements pertinents (constituants prosodiques) et extraire une pulsation (*via* les proéminences) qui lui permettront à la fois d'anticiper la suite de l'énoncé mais aussi de se rappeler du début de l'énoncé. L'auditeur doit ainsi s'adapter constamment à la parole perçue et donc au rythme de son interlocuteur, et ce de manière quasi instantanée (Peelle & Davis, 2012). L'interlocuteur doit également veiller à organiser son énoncé de manière à faciliter son décodage par l'auditeur. Dans cette section, nous verrons ainsi en quoi les caractéristiques rythmiques temporelles de la parole interviennent pour que son traitement puisse se réaliser en direct (*online*) efficacement.

Nous commencerons par présenter les mécanismes généraux mis en jeu lors de la synchronisation des rythmes attentionnels internes sur la pulsation d'un rythme externe avec la Théorie de l'Attention Dynamique (cf. 2.2.1). Dans un second temps, nous présenterons l'implication de la motricité dans la perception de la pulsation (cf. 2.2.2, p. 52). Pour finir, nous nous recentrerons sur les mécanismes de synchronisation attentionnelle appliqués plus spécifiquement à la parole, qui participent au décodage et à l'encodage du flux de parole dans le temps (cf. 2.2.3, p. 55).

2.2.1 *LA THÉORIE DE L'ATTENTION DYNAMIQUE*

Large et Jones (1999) suggèrent que la perception rythmique serait intimement liée à l'attention et proposent une théorie de l'attention appliquée aux événements rythmiques se déroulant dans le temps, la *Dynamic Attending Theory* (DAT³⁰). Ces auteurs observent que les théories de la structure temporelle ne permettent pas de rendre compte de la sensation rythmique qui se dégage de certains événements temporels naturels dont la structure n'est pas strictement régulière, comme la musique ou la parole. En effet, ces structures temporelles rythmiques

³⁰ Théorie de l'Attention Dynamique

présentent des variations de tempo ou de structure temporelle et pourtant, l'auditeur perçoit bien un rythme ou un mètre implicite, sous-jacent. De plus, ces variations temporelles dans la musique ou la parole sont souvent significatives, comme par exemple un ralentissement de tempo indiquant la fin d'une phrase (cf. 1.2.1.3.2, p. 21).

Selon Large et Jones (1999), les théories employées antérieurement à la DAT pour expliquer la représentation des structures temporelles des événements sont des théories de la mémoire : la perception de variations temporelles se fait par comparaison à une unité de mesure de référence encodée en mémoire. Ainsi, les modélisations statistiques de la structure temporelle proposent la construction d'un intervalle moyen comme trace mnésique de référence, tandis que d'autres théories de la structure temporelle proposent un encodage mnésique basé sur une pulsation de référence, à l'instar du modèle de l'horloge interne de Povel et Essens (1985). Si ces modèles sont utiles pour expliquer les représentations des structures temporelles, ils trouvent cependant leurs limites face aux variations temporelles des événements rythmiques (Large & Jones, 1999).

L'hypothèse de la perception d'une cohérence rythmique malgré des variations temporelles va alors amener Large et Jones (1999) à se rapprocher des théories du contrôle attentionnel, celui-ci impliquant la combinaison de deux processus attentionnels : 1/ un processus *top-down* orienté vers un but, reflétant les attentes actives et la focalisation de l'attention, et 2/ un processus *bottom-up*, la capture attentionnelle, reflétant les déplacements de la focalisation attentionnelle causés par des propriétés saillantes ou un changement inattendu dans la structure temporelle d'un événement externe. Les fluctuations de l'attention correspondraient ainsi aux oscillations internes, auxquelles les auteurs réfèrent sous le terme de « rythmes attentionnels ». Plus précisément, les rythmes attentionnels sont caractérisés par deux aspects : une oscillation périodique simple autonome, qui engendre des attentes temporelles et une pulsation d'énergie attentionnelle, qui indique le lieu et l'étendue temporelle de la focalisation attentionnelle. L'oscillation neuronale reflète les changements de polarité membranaire des neurones, alternant entre des états dépolarisés et hyperpolarisés. L'état dépolarisé correspond à une phase de haute excitabilité et l'état hyperpolarisé à une phase de basse excitabilité. La phase de haute excitabilité est la phase de sélection attentionnelle : l'information entrante qui coïncide temporellement avec la phase de haute excitabilité sera ainsi traitée plus efficacement (Peelle & Davis, 2012). Les rythmes attentionnels seraient de plus dynamiques, au sens où ils peuvent être entraînés par un rythme auditif externe qui ré-orient

la focalisation de l'attention et donc la perception, le décalage de phase des oscillations neuronales étant provoqué par les prééminences de l'événement extérieur.

Ce faisant, la DAT propose ainsi que les rythmes attentionnels (oscillations internes) génèrent des attentes permettant l'anticipation d'événements futurs (processus *top-down*) mais également qu'un événement rythmique externe entraîne les rythmes attentionnels, par la saillance ou prégnance de sa structure temporelle (processus *bottom-up*). Ce mécanisme d'entraînement résulterait d'une force d'attraction entre les deux rythmes, provoquant une mise en résonance neuronale et résultant en un couplage de phase et/ou de période de l'oscillation interne (Snyder & Large, 2005). Un rythme attentionnel pourrait ainsi ajuster en continu la phase et la période de ses fluctuations d'énergie attentionnelle sur les oscillations d'un rythme externe (Large & Jones, 1999).

Il faut cependant souligner que le mécanisme d'entraînement s'applique pour des rythmes simples caractérisés par une (pseudo) périodicité unique. Un rythme attentionnel interne s'y conforme bien, en revanche un rythme externe (musique ou parole par exemple) est généralement constitué de plusieurs périodicités (ou plusieurs niveaux de battement) entremêlées qui constituent sa structure métrique sous-jacente. Par exemple, en écoutant la chanson *Beat It* de Michael Jackson, l'auditeur peut choisir de suivre la pulsation sur chaque battement de batterie, ou un battement sur deux, ou deux fois plus vite soit sur chaque son de la basse, ou encore s'harmoniser avec plusieurs périodicités en synchronie, par exemple avec des effecteurs différents (pied, main, hanches, etc.). Cet enchevêtrement de périodicités externes (constitutives du mètre) implique plusieurs types de mécanismes d'entraînement : d'une part, un couplage uni-directionnel qui s'opère entre les périodicités externes (rythme complexe) et les rythmes attentionnels, chacune des différentes périodicités constitutives du rythme externe pouvant affecter un rythme attentionnel simple, mais pas réciproquement. D'autre part, un couplage bi-directionnel s'opère entre les rythmes attentionnels internes, pouvant s'influencer les uns les autres. Les relations entre les rythmes internes simples engagés dans l'entraînement à un rythme externe complexe font émerger par mimétisme oscillatoire la forme métrique hiérarchique sous-jacente du rythme externe (Large & Jones, 1999). Nous aborderons le traitement du mètre dans la parole plus en détail dans la partie 2.2.3.1 (p. 55).

2.2.2 *IMPLICATION DE LA MOTRICITE DANS LA PERCEPTION DE LA PULSATION*

Selon Patel et Iversen (2014), la perception d'une pulsation rythmique dans un signal auditif ne serait pas seulement liée à une synchronisation entre l'activité oscillatoire neuronale du cortex auditif et un signal auditif extérieur comportant des proéminences régulières (un morceau de musique par exemple), comme le propose la DAT, mais impliquerait un couplage bidirectionnel entre les aires cérébrales dédiées au traitement auditif et celles dédiées à la planification motrice. Patel et Iversen proposent ainsi l'hypothèse « *Action Simulation for Auditory Perception* » (ASAP) qui prévoit que les signaux neuronaux seraient activés depuis les régions auditives vers les régions de planification motrice pour renseigner ces dernières sur la temporalité des événements auditifs. Cette co-activation influencerait la temporalité des signaux périodiques de planification motrice en simulant ainsi un mouvement périodique qui engendrerait un signal neuronal en direction du système auditif, permettant de prédire quand tomberont les prochains battements de pulsation. Les battements de pulsation peuvent ensuite être intégrés dans des patrons hiérarchiques intégrant plusieurs niveaux de périodicités, les mètres. Dans ce cas, les différents niveaux de périodicités seraient associés à différents niveaux de prédiction et à la simulation de différents effecteurs de mouvements.

La perception de la pulsation n'est donc pas qu'un entraînement passif des réponses neuronales à un stimulus sonore, mais est un processus actif, volontaire et associatif entre audition et planification motrice. La perception d'une pulsation serait ainsi une forme de synchronisation invisible (Repp, 2005). La pulsation renvoie à une référence temporelle interne qui va façonner les rythmes perçus, guider l'attention dans le temps, influencer la perception des accents et des groupements. Dans certaines musiques dont le rythme n'est pas marqué de manière évidente acoustiquement, les auditeurs peuvent ne pas localiser la pulsation au même endroit (jazz notamment). Ainsi, le signal acoustique ne détermine pas la pulsation, et la pulsation perçue (activement donc) influence les rythmes perçus. La perception de la pulsation serait ainsi une fonction cognitive complexe impliquant une communication étroite entre les régions auditives et de planification motrice. Ce couplage entre régions auditives et motrices est par ailleurs plus important chez les musiciens (Grahn & Brett, 2007).

L'argument en faveur d'une implication des régions motrices pour la perception de la pulsation s'est développé suite à différentes observations expérimentales. Premièrement, la perception d'une pulsation musicale s'accompagne de l'activation des régions motrices (cortex prémoteur moyen, ganglions de la base, aire motrice supplémentaire, cervelet – Chen et al., 2008 ; Grahn & Brett, 2007), même lorsque l'auditeur ne bouge pas. En second lieu, la motricité

contribuerait à améliorer la focalisation attentionnelle. Par exemple, la réalisation de mouvements en synchronie avec un stimulus auditif influencerait la structure rythmique perçue (Chemin et al., 2014 ; Phillips-silver & Trainor, 2005 ; Phillips-Silver & Trainor, 2007) ou faciliterait la perception de la pulsation, en particulier chez les non-musiciens (Su & Pöppel, 2012).

De manière générale, les mouvements corporels permettent d'orienter les organes sensoriels pour une perception optimale de l'environnement. En effet, on peut par exemple penser aux saccades oculaires en perception visuelle, à l'exploration d'une surface par le toucher ou d'un lieu en marchant, ou encore aux mouvements de la tête pour orienter le pavillon de l'oreille et mieux percevoir un signal sonore (Morillon et al., 2014). De manière plus spécifique, les mouvements rythmiques périodiques peuvent également participer à une meilleure sélection attentionnelle et ainsi améliorer la perception (Large & Jones, 1999).

Par exemple, dans leur étude de 2014, Morillon et al. ont cherché à évaluer la part de l'attention et la part de la motricité sur le gain sensoriel dans une tâche de jugement de hauteur de notes à travers trois expériences. Dans les trois expériences, les participants entendaient plusieurs séquences de notes en même temps qu'une pulsation de fond. Les notes tombant sur les pulsations étant les cibles et les notes à contre-temps des distracteurs. Les séquences étaient présentées dans deux conditions expérimentales : 1/ écoute seule et 2/ écoute avec engagement moteur rythmique du participant (pression du doigt sur un bouton en même temps que la pulsation de fond). Les participants devaient juger si la moyenne des hauteurs des notes cibles d'une séquence était plus aiguë ou plus grave qu'une note de référence donnée au début de la séquence (la moyenne des distracteurs étant nulle).

Dans la première expérience, dans les deux conditions expérimentales (1/ et 2/ ci-dessus), les participants parvenaient à une bonne performance de catégorisation de hauteur, indiquant qu'ils avaient réussi à se focaliser sur les notes cibles tout en inhibant les distracteurs, ce qui reflète un effet de l'attention sur la sélection sensorielle. Cet effet sur la performance avec une focalisation sur les cibles et inhibition des distracteurs était renforcé dans la condition expérimentale 2/, indiquant que l'engagement moteur améliore la sélection sensorielle perceptive.

Dans une deuxième expérience, ces auteurs ont cherché à voir si les performances des participants étaient liées à la périodicité des stimuli et donc à leur caractère prédictible. Ils ont ainsi manipulé la périodicité des notes cibles et des distracteurs en introduisant un léger

décalage temporel (*jitter*), menant à 4 sous-conditions expérimentales : une sous-condition rythmique isochrone (sans décalage), deux sous-conditions semi-rythmiques (décalages touchant les cibles ou les distracteurs) et une sous-condition arythmique (décalage touchant les cibles et les distracteurs) ; le reste de l'expérience étant identique à la première expérience. Les résultats ont été cohérents avec ceux de la première expérience, renforçant l'hypothèse de la part de l'attention et l'hypothèse de la part additionnelle de la motricité dans le processus de sélection perceptive. En revanche, le décalage temporel a provoqué une dégradation de l'effet mélioratif de la motricité, indiquant que la contribution de la motricité serait dépendante du caractère prévisible des cibles.

Dans une troisième expérience enfin, les auteurs ont cherché à savoir si l'effet bénéfique additionnel de la motricité sur la sélection sensorielle était dépendant de sa co-occurrence avec les fluctuations temporelles de l'attention. Ils ont ainsi échangé les rôles des cibles et des distracteurs : les participants devaient cette fois-ci porter leur attention sur les notes à contre-temps des pulsations de fond (devenues les cibles) et inhiber les notes tombant sur les temps (devenues les distracteurs). Dans la condition expérimentale 2/ (avec engagement moteur rythmique), les participants pressent le bouton sur les pulsations, renvoyant cette fois-ci aux distracteurs. Les fluctuations de l'attention et de la motricité sont donc en opposition de phase. Trois hypothèses sont formulées :

- H1/ Avantage sélectif pour la motricité : toute activité motrice dirige et renforce la sélection sensorielle. La comparaison de performance entre les conditions expérimentales 1/ et 2/ devrait alors montrer cette fois-ci une baisse de performance en condition 2/ pour la sensibilité aux cibles (notes à contre-temps), puisque l'activité motrice est associée dans cette troisième expérience aux distracteurs (notes sur le temps).
- H2/ Avantage sélectif pour l'attention : l'activité motrice renforce l'effet de l'attention sur la sélection sensorielle. La force de l'effet entre les conditions 1/ et 2/ devrait être identique, que ce soit dans l'expérience 1 ou dans l'expérience 3.
- H3/ Avantage sélectif pour une synergie entre motricité et attention : l'activité motrice et la sélection attentionnelle doivent être co-occurentes pour qu'un effet bénéfique survienne. Il ne devrait pas y avoir de différence entre les conditions expérimentales 1/ et 2/ dans la troisième expérience puisque les fluctuations de la motricité et de l'attention ne sont pas en phase.

Les résultats de cette troisième expérience ne montrent pas de différence significative entre les deux conditions expérimentales, allant alors dans le sens de l'hypothèse H3 et indiquant donc la nécessité d'une co-occurrence entre les fluctuations de la motricité et de l'attention pour obtenir un effet bénéfique additif de la motricité sur la sélection attentionnelle sensorielle.

Les résultats de ces trois expériences sont précieux au sens où ils nous éclairent sur l'influence *top-down* de l'attention comme processus crucial actif de sélection sensorielle auditive rythmique ; sur l'importance de la périodicité du stimulus sensoriel pour bénéficier de l'effet additionnel de la motricité sur l'attention ; et enfin sur l'influence *top-down* de la motricité sur l'attention et de fait sur la sélection sensorielle lorsque leurs fluctuations sont en phase, ce qui permet de faire des prédictions sur la survenue de nouveaux événements sensoriels, facilitant leur traitement.

2.2.3 SYNCHRONISATION ET ENTRAINEMENT DANS LA PAROLE

2.2.3.1 COUPLAGE DES PERIODICITES PROSODIQUES ET ATTENTIONNELLES

La planification phonétique de la parole, et notamment son organisation temporelle, peut être affectée par différents facteurs directs et indirects. La structure prosodique, le type de réalisation vocalique ou consonantique, le type des segments adjacents ainsi que d'autres facteurs extra-linguistiques tels que la vitesse ou le style de parole ont un effet direct sur la temporalité. D'autres facteurs linguistiques participent à la prédictibilité de l'énoncé et vont avoir ainsi un effet sur la structure prosodique, et par conséquent sur la temporalité de la parole de manière indirecte : la syntaxe, la longueur de l'énoncé, la sémantique, la pragmatique et le lexique. Selon l'hypothèse de la redondance harmonieuse du signal (*Smooth signal redundancy hypothesis*), la structure prosodique participe à rendre équivalente la probabilité de reconnaissance de chaque élément du signal de parole en tenant compte de tous les facteurs extra-linguistiques et linguistiques. Par exemple, la saillance acoustique d'un mot peu prévisible sera renforcée (proéminence et/ou frontières prosodiques renforcées) par rapport à un mot attendu (Turk & Shattuck-Hufnagel, 2014). Ainsi, la temporalité, *via* la prosodie et donc le rythme, participe à la compréhension de la parole.

En 1990, Pitt et Samuel proposent l'*Attentional Bounce Hypothesis*³¹ (ABH) afin de questionner la part du rythme dans le traitement de la parole. L'attention serait répartie de manière différentielle sur les segments du flux de parole continue, avec une allocation plus importante pour les syllabes accentuées qu'inaccentuées. L'apparition quasi régulière de syllabes accentuées au milieu de syllabes inaccentuées formerait une sorte de grille métrique à laquelle l'attention de l'auditeur se synchroniserait pour « rebondir » (*bounce*) de syllabe accentuée en syllabe accentuée et anticiper les syllabes accentuées à venir. Le temps de traitement et le gain perceptif seraient ainsi améliorés sur les syllabes accentuées. Les auteurs explorent leur hypothèse au moyen de trois expériences présentant des tâches de détection phonémique (temps de réaction) au sein de phrases ou de listes de mots bisyllabiques présentant ou non un même patron métrique. Ils en concluent que les auditeurs forment des attentes rythmiques induites par l'alternance des syllabes accentuées et inaccentuées, avec une allocation de ressources attentionnelles en faveur des syllabes accentuées. Ce faisant, cette différence d'allocation attentionnelle est d'autant plus importante que la structure métrique est régulière, renforçant la prévisibilité accentuelle et donc la direction de l'attention.

De même que nous l'avons vu avec la DAT (cf. 2.2.1, p. 49), Peelle et Davis (2012) proposent que pour s'adapter au rythme de la parole entendue, les oscillations neuronales opèrent un verrouillage ou couplage de phase : un système quasi périodique (ici, les oscillations neuronales) s'ajuste pour se synchroniser (osciller en même temps) à un stimulus externe quasi périodique (ici, le rythme de la parole). Le système oscillatoire de l'auditeur peut donc à la fois décaler sa phase et adapter sa fréquence d'oscillation à celle du signal de parole perçu. De cette manière, les oscillations neuronales alignent leurs phases de haute excitabilité avec les éléments saillants de la parole perçue afin de les traiter plus efficacement et de prédire l'arrivée des nouveaux éléments pertinents. De plus, le processus de verrouillage de phase ne dépend pas uniquement des informations acoustiques de la parole : afin d'ajuster leur période, les oscillations neuronales sont également dépendantes de l'attention dirigée par l'auditeur envers l'information entrante.

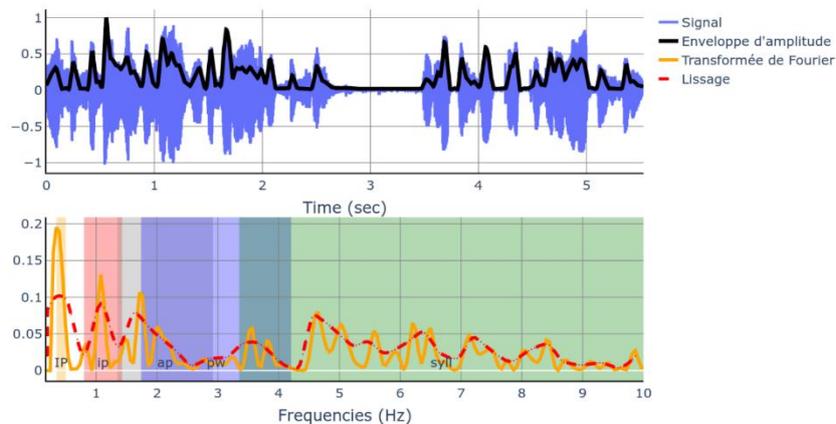
D'autre part, dans la parole, l'entraînement neuronal de différentes bandes de fréquence reflète les modulations hiérarchiques dans l'enveloppe d'amplitude de la parole (Aichert et al., 2021). L'enveloppe d'amplitude correspond à la somme des puissances acoustiques des fréquences du signal de parole en fonction du temps. La quasi-régularité de la parole se trouve

³¹ Hypothèse du rebond attentionnel.

dans les modulations d'amplitude de basse fréquence du signal sonore (4 – 8 Hz). Elle est liée au cycle d'ouverture et fermeture de la mâchoire ainsi qu'à la production de voisement et est associée au niveau syllabique. En d'autres termes, la puissance acoustique sera généralement plus importante (et donc saillante) lorsque la bouche est grande ouverte pour la production d'un son voisé, par exemple le son /a/, que dans l'intervalle entre deux syllabes (Pelle & Davis, 2012). La structure rythmique de la parole permet d'entraîner la coordination des mouvements articulatoires ainsi que la perception auditive, posant une ossature sur laquelle peuvent s'organiser d'autres caractéristiques temporelles de la parole (Pelle & Davis, 2012). Le verrouillage de phase qui opère lors de l'écoute de parole pourrait créer des attentes rythmiques facilitant les productions ultérieures, tout comme dans la régulation de l'alternance des tours de parole (Aichert et al., 2021). La transformée de Fourier de l'enveloppe d'amplitude de la parole révèle plusieurs pics d'amplitude renvoyant aux périodicités associées aux différents niveaux prosodiques (Vaysse, 2023). A partir de l'exemple en **Figure 12**, on remarque ainsi différents pics dans le lissage de la transformée (courbe rouge) : un pic correspondant au niveau du syntagme intonatif aux alentours de 0,3 Hz (3,3 s.), un pic au niveau du syntagme intermédiaire vers 1,1 Hz (909 ms), un pic au niveau du syntagme accentuel/mot prosodique vers 1,6 Hz (625 ms). Ces périodicités dans le signal de parole (rythme externe) peuvent être suivies par les rythmes attentionnels internes de différentes bandes de fréquences : les indices prosodiques sont suivis par les oscillateurs delta (0,5 à 2 Hz), le niveau syllabique par les oscillateurs thêta (4 à 8 Hz) et les informations segmentales par les oscillateurs bêta et gamma (25 à 35 Hz) (pour une revue, cf. Te Rietmolen, 2019). Lorsque différents niveaux de périodicités internes sont entraînés conjointement sur différents niveaux de constituance ou de prééminence d'un signal de parole rythmique externe, ces périodicités vont être hiérarchisées, laissant transparaître la structure métrique sous-jacente.

Figure 12 :

Transformée de Fourier de l'enveloppe d'amplitude d'un signal de parole



Note. Partie supérieure : en bleu : variations d'amplitude du signal de parole au cours du temps ; en noir : enveloppe d'amplitude du signal. Partie inférieure : courbe jaune : transformée de Fourier à partir des périodicités liées aux pics d'amplitude du signal de parole ; courbe rouge : lissage de la transformée de Fourier ; les zones colorées de gauche à droite renvoient aux niveaux prosodiques (en orange : syntagme intonatif (IP) ; en rouge : syntagme intermédiaire (ip) ; en gris : syntagme accentuel (ap) ; en bleu : mot prosodique (pw) ; en vert : syllabe (syll). Tiré de Vaysse (2023, p. 97).

2.2.3.2 ALTERATION DU COUPLAGE

Lorsque le verrouillage de phase est altéré, la compréhension du signal de parole l'est également. Au contraire, lorsque la parole est intelligible, le verrouillage de phase est amplifié dans l'hémisphère gauche. Cependant, le verrouillage de phase n'implique pas que le signal soit compréhensible : le mécanisme de verrouillage serait plutôt une étape préalable nécessaire au traitement général des informations sensorielles, et permettrait de mener à un signal intelligible si le verrouillage est amplifié dans l'hémisphère gauche, reflétant un traitement d'informations linguistiques. Ainsi, le traitement linguistique renforce le couplage de phase et participe à faciliter le traitement et la prédiction des indices acoustiques (Pelle & Davis, 2012).

Une altération du couplage de phase en perception de parole peut par exemple avoir lieu lorsqu'une prééminence apparaît à un moment inattendu ou encore qu'un segment de l'énoncé perçu est accéléré ou ralenti par rapport au reste de la séquence.

En manipulant le paramètre de durée pour provoquer un déplacement de l'accent à l'intérieur du mot, Magne et al. (2007) ont montré qu'une incongruité métrique pouvait se répercuter sur le traitement lexical des mots et notamment avoir une incidence sur leur traitement sémantique. Pour ce faire, ils ont procédé au moyen de deux tâches expérimentales d'écoute de courtes phrases en français : une tâche de jugement métrique et une tâche de jugement sémantique, avec recueil de données comportementales et électroencéphalographiques (EEG). Dans chacune des deux tâches, le dernier mot était ou non

manipulé pour provoquer une incongruité sémantique (mot inattendu) et/ou métrique (structure métrique inattendue due au déplacement de l'accent). Ainsi, dans la tâche de jugement sémantique, les auteurs ont manipulé à la fois l'incongruité sémantique et l'incongruité métrique, et réciproquement pour la tâche de jugement métrique. Dans la tâche de jugement sémantique, les auteurs ont observé que même lorsqu'il n'y avait pas d'incongruité sémantique, le déplacement de l'accent (provoquant une incongruité métrique) était associé à l'élicitation d'une composante N400, celle-ci étant normalement associée à une incongruité sémantique. Selon les auteurs, ce résultat indiquerait à la fois un potentiel traitement automatique de l'information métrique ainsi qu'une implication de la structure métrique dans l'accès au sens. Ces résultats sont de plus confortés par les données comportementales recueillies dans la tâche de jugement sémantique en condition de congruité sémantique (mot attendu), montrant un taux d'erreur et un temps de réaction plus élevés en présence d'incongruités métriques (Magne et al., 2007).

De plus, les modulations d'amplitude participent également à la catégorisation phonémique : les indices rythmiques de l'enveloppe, telle que la vitesse d'élocution, peuvent influencer la perception des frontières catégorielles entre phonèmes. Une altération de la synchronisation ou un changement abrupt de la vitesse d'élocution peuvent ainsi provoquer un changement de catégorisation et donc altérer la compréhension du signal. Peelle et Davis (2012) rapportent ainsi les résultats de plusieurs études (Port, 1979 ; Miller, 1981 ; Summerfield, 1981 ; Miller et al., 1984, 1997) lors desquelles les participants devaient porter un jugement sur la syllabe perçue : « /pa/ », « /ba/ » ou une troisième catégorie phonémique intermédiaire ambiguë, « /?a/ », ces trois catégories (/pa/, /ba/ et /?a/) variant selon leur délai de voisement. Lorsque la parole est accélérée, le délai de voisement de /?a/ se retrouve perçu plus long et le phonème ambigu /?/ est donc assimilé à la catégorie /p/ ; tandis que dans la parole ralentie, le délai de voisement de /?/ se retrouve perçu plus court et /?/ est assimilé à la catégorie /b/. Peelle et Davis (2012) proposent comme explication que lorsque les oscillations cérébrales sont synchronisées sur les oscillations du rythme de la parole perçue à un rythme naturel, si la consonne arrive au niveau d'une région de basse excitabilité, la syllabe sera perçue comme /pa/ (délai de voisement plus allongé), mais si la consonne arrive dans une région de haute excitabilité, la syllabe sera perçue comme /ba/ (délai de voisement plus court). Lorsque la parole est accélérée, le phonème ambigu perçu /?/ va se retrouver en zone de basse excitabilité, le faisant passer dans la catégorie /pa/ ; tandis que dans la parole ralentie, /?a/ se retrouve en zone

de haute excitabilité, et est perçu /ba/. En résumé, le traitement du délai de voisement varie selon la vitesse de la parole.

Une altération du couplage et donc de la compréhension peut également survenir dans certains troubles du langage, dans des conditions d'écoute dégradée ou encore lors de l'exposition à une langue étrangère. Par exemple, le trouble neurodéveloppemental du langage (TNL) peut se manifester par une altération de la production et de la compréhension de la parole, qui se traduit par une plus grande difficulté à suivre et détecter les éléments rythmiques pertinents dans le flux de parole et donc à segmenter et décoder la parole entrante (Corriveau et al., 2007). Guiraud et al. (2018) rapportent ainsi une tendance à une plus grande difficulté des enfants TNL par rapport à un groupe contrôle à détecter des incongruïtés sémantiques dans la parole perçue, cette différence de sensibilité devenant significative en cas de perception de parole accélérée ou compressée. De même, Port et al. (1995) rapportent une intelligibilité plus faible de la parole présentant un accent étranger. Lorsque le rythme de la parole est altéré dans le cas d'un trouble moteur de la production de parole comme le bégaiement (Monfrais-Pfauwadel & Bijleveld, 2016), l'aphasie non fluente (Baqué et al., 2016), l'apraxie ou encore la dysarthrie (Aichert et al., 2021), nous pouvons supposer que l'interlocuteur aura plus de mal à se synchroniser à l'énoncé produit, affectant l'interaction et la communication de manière plus globale.

Pour résumer, nous retenons que la synchronisation des rythmes attentionnels sur la pulsation rythmique d'un signal de parole sous-tend les mécanismes cognitifs spécifiques au traitement de la parole, en permettant à l'auditeur de décoder (segmenter) le signal au fur et à mesure qu'il se déroule dans le temps mais également d'anticiper la suite de l'énoncé et donc de faciliter son traitement, grâce au suivi de la pulsation (succession des proéminences). La synchronisation est dans ce cas permise lorsque l'énoncé produit par le locuteur présente des repères pertinents et prévisibles sur lesquels l'attention de l'auditeur peut rebondir et être entraînée. De plus, la synchronisation faciliterait également la production de parole subséquente, le patron métrique (les futures pulsations) étant déjà posé : les attentes générées par la perception sont réalisées dans la production (Aichert et al., 2021). L'implication de la motricité (geste co-occurent avec la pulsation) participerait à faciliter l'extraction de la pulsation, la synchronisation et la génération des attentes et donc à améliorer la perception et la production.

2.3 LANGAGE ET SUPRAMODALITE DU RYTHME

Dans la partie précédente (2.2.3, p. 55), nous avons vu l'importance de l'établissement d'une synchronisation entre les rythmes attentionnels et les indices rythmiques de la parole pour traiter efficacement le flux de parole. Le processus de synchronisation neuronale reposant sur des mécanismes généraux du traitement attentionnel (cf. 2.2.1, p. 49 et 2.2.2, p. 52), nous pouvons donc nous poser la question suivante : existe-t-il une supramodalité du rythme, qui dépasserait et sous-tendrait les modalités spécifiques (parole, musique, motricité, etc.), comme le suggère Astésano (2022) ? Nous présentons des éléments de discussion en partant d'abord de l'existence de déficits communs entre traitement du rythme et traitement du langage (cf. 2.3.1). Dans un second temps, nous évoquons les similitudes entre traitements de la musique et du langage (cf. 2.3.2, p. 64). Nous terminons enfin avec la présentation d'études sur l'amorçage rythmique de la parole qui nous intéressent tout particulièrement (cf. 2.3.3, p. 67).

2.3.1 TROUBLES DU LANGAGE ET TROUBLES DU RYTHME

Le traitement du rythme auditif jouant un rôle essentiel dans le traitement du signal de parole, et notamment dans le traitement de l'enveloppe temporelle, un déficit général de traitement du rythme pourrait se répercuter sur le traitement langagier. En effet, plusieurs travaux rapportent des liens entre les capacités générales de traitement du rythme et les capacités de traitement du langage.

Selon la démarche neuropsycholinguistique (Nespoulous, 2016), la description et la mise en correspondance des manifestations de surface avec des déficits sous-jacents permet de construire l'architecture fonctionnelle du langage et des processus de traitements engagés. L'étude des troubles de la parole et du langage nous permet ainsi d'en apprendre plus sur les mécanismes en jeu dans le traitement du langage (Nespoulous, 2016 ; Sahraoui, 2015). En partant de l'observation de déficits affectant de façon potentiellement concomitante les mécanismes de traitement du rythme et du langage, la démarche neuropsycholinguistique permet de mieux situer le rôle du rythme dans le traitement du langage et soutient l'hypothèse de mécanismes partagés ou de recouvrements entre cognition du rythme et cognition langagière.

Afin de mesurer les capacités de traitement du rythme, et de les considérer ensuite en lien avec les capacités préservées ou déficitaires de traitement du langage, de nombreuses tâches et batteries ont été développées. Parmi les plus utilisées, nous pouvons citer les tâches de Tempo

Moteur Spontané (TMS) et de Synchronisation Sensori-Motrice (SMS³²). Le TMS consiste en un mouvement de frappe (*unpaced tapping*) avec le doigt ou la main effectué à une vitesse naturelle et agréable et reflète la période de référence de l'individu (Fraisse, 1974) (cf. 1.2.1.3.1, p. 19). La Synchronisation Sensori-Motrice (SMS) consiste à exécuter un mouvement du doigt ou de la main en rythme (en synchronie) avec la période d'un stimulus rythmique externe (*paced tapping*) et reflète du point de vue de la Théorie de l'attention dynamique (DAT) l'adaptation dynamique de la période de référence à un stimulus externe (cf. 2.2.1, p. 49). La synchronisation se distingue du mécanisme de réaction à un stimulus et repose au contraire sur un mécanisme d'anticipation de la pulsation (ou battement) à venir, l'anticipation étant rendue possible par le caractère régulier des pulsations dans le stimulus externe rythmique (Fraisse, 1974). Des batteries de tests incluant les tâches de TMS et SMS parmi d'autres épreuves permettent d'évaluer les capacités de traitement du rythme de manière exhaustive, comme par exemple la batterie BAASTA (*Battery for the Assessment of Auditory Sensorimotor Timing Abilities*) de Dalla Bella et al. (2017), qui est constituée d'un ensemble de huit tâches de perception et production (spontanée et synchronisation sensorimotrice) rythmiques et a été étalonnée auprès d'adultes musiciens et non-musiciens.

L'existence de troubles du traitement du rythme a pu être observée de manière cooccurrence à différents déficits langagiers, nous pouvons relever par exemple :

- des difficultés de traitement en perception rythmique (tâche de discrimination) en lien avec un déficit de conscience phonologique chez des enfants d'âge préscolaire et scolaire (Corriveau et al., 2007 ; Ozernov-Palchik et al., 2018) ;
- des difficultés en production rythmique (tâche de synchronisation-continuation) cooccurrentes avec un trouble du langage écrit chez des enfants d'âge scolaire (Corriveau & Goswami, 2009 ; Lê et al., 2020 ; A. T. Tierney & Kraus, 2013) ;
- l'existence d'associations entre les capacités de traitement en perception et production rythmique et les capacités de traitement du langage écrit chez l'adulte (neurotypique ou avec dyslexie développementale) (Thomson et al., 2006) ;
- une plus grande difficulté de SMS (extraits musicaux ou métronome) chez des enfants et adolescents qui bégaièrent par rapport à un groupe contrôle (Falk et al., 2015) ;

³² De l'anglais *Sensori-Motor Synchronization*.

- une plus grande asynchronie dans une tâche de SMS (mélodies simples à jouer sur un piano digital en rythme avec un métronome) chez un groupe d'adultes présentant un bégaiement par rapport à un groupe contrôle (van de Vorst & Gracco, 2017).

De même, dans leur revue théorique sur le partage de mécanismes entre le traitement du rythme dans la parole et la musique (PRISM³³), Fiveash et al. (2021) relatent également le partage de déficits communs : les autrices se concentrent dans leur revue sur trois troubles du langage, la dyslexie développementale (trouble du langage écrit et de la conscience phonologique), le trouble développemental du langage (trouble de l'acquisition du langage oral) et le bégaiement (trouble de la fluence de la parole). La dyslexie et le trouble neurodéveloppemental du langage, potentiellement co-morbides, présentent tous deux une atteinte du traitement auditif précis et de la synchronisation des oscillations neuronales à un rythme externe. Cela se traduit sur le versant du rythme musical par des difficultés dans la production, perception et synchronisation à un rythme externe, et sur le versant du rythme langagier par un déficit dans le traitement de l'attaque syllabique, de la perception accentuelle et plus largement de l'enveloppe temporelle de la parole. Concernant le bégaiement, les autrices rapportent des atteintes du traitement de la perception, de la production et de la synchronisation à un rythme externe ainsi que des atteintes de la fluence de la parole (*patterns* temporels irréguliers).

Ladányi et al. (2020) proposent ainsi l'hypothèse *Atypical Rhythm Risk Hypothesis*³⁴ (ARRH), selon laquelle un trouble du traitement du rythme serait un potentiel facteur de risque prédisant un trouble du langage. Pour cela, les autrices s'appuient sur l'observation (1) d'associations entre capacités de traitement du rythme et performances langagières chez l'enfant et l'adulte au développement typique, (2) d'associations entre les mécanismes sous-jacents au traitement du rythme et aux traitements de la perception et de la production du langage (traitement temporel fin, synchronisation neuronale, implication des fonctions motrices, traitement hiérarchique) et (3) de la prévalence de troubles du traitement du rythme chez les enfants au développement langagier atypique (TNL). Elles argumentent ainsi en faveur d'un dépistage précoce dans l'enfance des troubles du traitement du rythme au moyen d'épreuves rythmiques musicales, qui permettrait d'orienter ensuite les patients éventuels pour un bilan orthophonique. Cette proposition rejoint les résultats préliminaires de Daigmorte et al. (2022) à propos de l'utilisation de méthodes rythmiques pour la rééducation du langage, qui

³³ Approche *Processing Rhythm In Speech and Music*.

³⁴ Hypothèse de risque de rythme atypique.

indiquent que les orthophonistes considèrent le rythme comme un prérequis pour le développement du langage mais manquent d'arguments scientifiques et d'outils pour davantage intégrer l'évaluation et la rééducation des capacités de traitement du rythme dans les protocoles d'évaluation et de rééducation de la parole et du langage.

En raison de déficits de traitement communs au rythme musical et au rythme langagier soutenant l'hypothèse de processus sous-jacents partagés, Fiveash et al. (2021) suggèrent qu'exercer le rythme musical pourrait améliorer à la fois le traitement de la musique et de la parole. Dans la partie suivante sur les liens entre langage et musique (2.3.2, p. 64), nous verrons ainsi d'une part dans quelle mesure les mécanismes de traitement sont partagés entre les deux domaines (2.3.2.1, p. 64), et d'autre part l'effet de la pratique musicale sur le traitement du langage (2.3.2.2, p. 65).

2.3.2 LANGAGE ET MUSIQUE

2.3.2.1 MECANISMES DE TRAITEMENT PARTAGES

La musique, que Brandt et al. (2012) définissent comme un 'jeu créatif avec les sons', occupe une place importante chez l'Homme ; elle traverse les cultures, les âges et jouerait un rôle dans le développement cognitif général et l'acquisition du langage. En raison du partage structurel de la musique et du langage, la frontière entre les deux domaines n'est d'ailleurs pas toujours nette : comment catégoriser objectivement par exemple le chant, le slam, le rap, les langues sifflées ou tambourinées ? Le traitement différentiel de la musique et du langage se ferait graduellement au cours du développement, à partir des expériences culturelles. Chez le nouveau-né, le langage oral serait ainsi traité comme une forme de musique, *via* la prosodie de la parole : les sons sont discriminés selon leurs contrastes structurels acoustiques puis progressivement associés aux aspects syntaxiques et sémantiques du langage. Avant six mois, l'enfant peut ainsi discriminer des sons venant de n'importe quel système linguistique, avant que sa perception ne devienne plus spécifique. Cette spécialisation linguistique opérée par le nouveau-né est facilitée par la mise en évidence des contrastes prosodiques dans le langage adressé à l'enfant (*motherese*) ou les comptines, qui vont exagérer les contours mélodiques et rythmiques de la parole, renforçant l'attention de l'enfant sur les structures prégnantes de la langue (Dodane et al., 2018). Le traitement des propriétés musicales de la prosodie de la parole étant nécessaires pour le décodage du flux sonore, l'entraînement explicite de ces propriétés par

la pratique musicale devrait se manifester par une amélioration du traitement du signal de parole.

2.3.2.2 EFFET DE LA PRATIQUE MUSICALE SUR LE TRAITEMENT DU LANGAGE

D'après Patel (2011), les avantages attribués aux musiciens dans le traitement des sons de la parole auraient deux causes possibles : (1) une origine génétique, selon laquelle un individu ayant une ouïe plus fine sera plus susceptible de suivre un parcours musical et de mieux percevoir les sons de la parole, et (2) une origine liée à la pratique musicale, laquelle entraînerait des modifications cérébrales fonctionnelles et structurelles et donc un meilleur traitement de la parole.

L'idée d'un transfert de compétences entre la pratique musicale et le traitement du langage a notamment été étudiée par Patel (2011, 2014), qui propose l'hypothèse OPERA³⁵ : la précision, l'émotion, la répétition et l'attention associées à la pratique musicale stimuleraient les corrélats neuronaux communs au traitement de la musique et du langage, provoquant des modifications cérébrales fonctionnelles et structurelles et résultant en un bénéfice pour le traitement musical mais également pour le traitement langagier. Plus concrètement, selon Moussard et al. (2012) la pratique musicale entraînerait une modification des aires corticales impliquées dans la séquentialité, la simultanéité, la précision gestuelle et métrique et la mémoire. En effet, des différences de traitement du langage se retrouvent selon le degré de pratique musicale (Besson et al., 2011), concernant par exemple la segmentation de la parole (François & Schön, 2011), la discrimination des sons de la parole (Parbery-Clark et al., 2009, 2012), le traitement de la hauteur (Schön et al., 2004) ou encore la précision de l'articulation (Cason et al., 2020). La sensibilité au rythme aiderait à suivre le déroulé de parole dans des conditions d'écoute dégradées, il y aurait ainsi un avantage du musicien dans la perception de parole dans le bruit (Parbery-Clark et al., 2009, 2012 ; Slater & Kraus, 2016).

Lorsque les mécanismes sous-jacents au traitement du rythme de la parole sont surentraînés par la pratique musicale, cela se traduirait par des améliorations dans le traitement du langage. En effet, des avantages résultant de l'apprentissage de la musique sur le traitement du langage ont été mis en évidence chez des participants au développement typique ainsi qu'auprès de participants présentant des troubles du langage.

³⁵ *Overlap* (recouvrements), *Precision*, *Emotion*, *Repetition*, *Attention*.

Par exemple, Moritz et al. (2013) ont montré que des enfants d'âge préscolaire ayant reçu une formation musicale intensive sur l'année s'améliorent à la fois dans les tâches rythmiques (discrimination et SMS) et les tâches de conscience phonologique (les tâches rythmiques et de conscience phonologique étant significativement corrélées) par rapport aux enfants ayant reçu une formation musicale moins poussée. Zanto et al. (2024) rapportent des résultats similaires (corrélations entre les habiletés rythmiques et les capacités de traitement du langage écrit) avec un entraînement longitudinal basé uniquement sur le rythme (application ludique sur tablette) chez des enfants d'âge scolaire. La pratique musicale s'accompagne également d'améliorations significatives concernant le traitement temporel de la parole (durée syllabique et délai de voisement), le développement des représentations phonologiques et la segmentation de la parole par rapport à des cours de peinture, chez des enfants au développement typique (Chobert et al., 2014 ; François et al., 2013). L'apprentissage de la musique peut également avoir un effet bénéfique sur le langage écrit et notamment la remédiation des troubles graphomoteurs (Véron-Delor et al., 2017).

Des effets de transferts positifs de la pratique musicale sur le traitement du rythme et du langage, et notamment sur la perception catégorielle et auditive, l'attention auditive, la conscience phonologique, la lecture et la répétition de pseudo-mots ont également été retrouvés auprès d'enfants dyslexiques (Flaunacco et al., 2015 ; Habib et al., 2016 ; Moussard et al., 2012). Dans le domaine de la rééducation de la parole, nous relevons par exemple la *Melodic Intonation Therapy* (MIT) (M. L. Albert et al., 1973), qui est une méthode de rééducation de l'aphasie basée sur la répétition de phrases chantées, en *sprechgesang* (chant parlé) ou parlées, guidées par le battement de la main. Ainsi, le caractère ancré et rythmique de la musique (par le jeu d'un instrument, le chant ou le battement de la pulsation) semble aider à initier et entraîner la parole. Selon Stahl et al. (2011), l'effet de transfert positif associé à la rééducation musicale de la MIT serait en fait lié à un effet de la composante rythmique, et ce en particulier pour les patients présentant des lésions des ganglions de la base, ce qui indiquerait leur rôle dans la segmentation rythmique de la parole. D'autres méthodes de rééducation basées sur l'indigage rythmique, qui consiste à accompagner la parole par une stimulation auditive rythmique (pulsation de métronome, motif rythmique ou extrait musical), s'avèrent efficaces auprès de patients présentant une dysarthrie, un bégaiement, une dyslexie ou une apraxie de la parole en menant à une amélioration de la fluidité, du débit, du rythme et de l'intelligibilité de la parole (Thaut & Hoemberg, 2019). De plus, l'écoute régulière de musique (et non la pratique) aurait un effet comparable à celui de l'écoute régulière de livres audio sur la mémoire auditive

sensorielle et donc sur la compréhension de la parole par rapport à une condition contrôle, notamment chez des patients cérébrolésés (Särkämö et al., 2010).

2.3.3 EFFET DE L'AMORÇAGE RYTHMIQUE SUR LE TRAITEMENT DU LANGAGE

Des bénéfices à court terme (effets immédiats) sur le traitement du langage suite à une stimulation musicale ou rythmique peuvent également survenir. Selon les propositions de la *Dynamic Attending Theory* (DAT, cf. 2.2.1), l'exposition à un stimulus rythmique auditif, aussi appelée « amorçage rythmique », induirait des attentes temporelles sur la perception d'un stimulus auditif subséquent. Cet effet a pu être testé sur le versant de la perception et de la production de la parole, auprès de populations neurotypiques ou présentant des troubles du langage de tous âges et de différentes langues.

Nous avons pu recenser les différents effets d'amorçage rythmique sur le langage suivants :

- l'effet d'un amorçage musical (extraits musicaux, sons de percussions ou tonalités) ou langagier ;
- l'effet de la régularité de l'amorce ;
- l'effet de la similarité (congruence) métrique entre la structure métrique de l'amorce et celle de la cible langagière ;
- l'effet d'un amorçage auditif seul ou audio-moteur (amorce auditive accompagnée d'un engagement moteur du participant).

Concernant le versant de la perception de la parole, la présentation d'un stimulus musical rythmique régulier se traduit par de meilleures capacités de traitement (jugement de grammaticalité) par rapport à un stimulus irrégulier ou une condition sans amorçage, chez des adultes neurotypiques, des patients parkinsoniens ou dyslexiques, des enfants au développement typique ou présentant un trouble spécifique du langage ou une dyslexie (Bedoin et al., 2016 ; Canette et al., 2019, 2020 ; Kotz & Gunter, 2015 ; Przybylski et al., 2013). Les capacités de traitement en perception phonologique sont meilleures après un amorçage musical rythmique dont la structure métrique est similaire (congruente) au syntagme cible par rapport à un amorçage rythmique métriquement incongruent, chez des adultes neurotypiques francophones (Cason, Astésano, et al., 2015 ; Cason & Schön, 2012). Cet effet de congruence

métrique est de plus renforcé lorsque les participants effectuent un court exercice moteur³⁶ rythmique lors de la passation expérimentale (Cason, Astésano, et al., 2015). D'autres études ont également pu montrer que les capacités perceptives d'un changement lexical sont meilleures après la présentation d'un amorçage musical rythmique audio-moteur³⁷ métriquement congruent par rapport à un amorçage audio-moteur incongruent, et sont meilleures après un amorçage rythmique audio-moteur (métriquement congruent ou incongruent) par rapport à un amorçage rythmique auditif ou une condition sans amorçage, chez des adultes neurotypiques germanophones et francophones (Falk et al., 2017 ; Falk & Dalla Bella, 2016).

Concernant à présent le versant de la production, la présentation d'une amorce rythmique musicale métriquement congruente se traduit par un temps d'initiation de la parole réduit ou une meilleure précision phonologique lors d'une tâche de lecture, par rapport à un amorçage irrégulier, métriquement incongruent ou une condition sans amorçage, chez des enfants sourds, adultes neurotypiques ou avec la maladie de Parkinson, une aphasie ou une apraxie de la parole, en témoignent les résultats obtenus auprès de populations germanophone (Aichert et al., 2021), francophone (Cason, Hidalgo, et al., 2015 ; Contreras Roa et al., 2022), anglophone (Gould et al., 2016) et sinophone (Zhang & Zhang, 2019). De plus, la présentation d'une amorce rythmique langagière régulière se traduit par un temps d'initiation de la parole réduit lors d'une tâche de lecture, par rapport à une amorce irrégulière, chez des participants neurotypiques germanophones (Späth et al., 2016). L'effet positif d'un amorçage rythmique régulier (métronome) sur la fluence de la parole auprès de participants présentant un bégaiement est observé notamment par Brady (1969), par rapport à une condition sans amorçage, à une tâche non rythmique faisant office de distraction ou à un métronome arythmique. Par ailleurs, Jungers et al. (2016) ont montré l'effet de la vitesse d'un amorçage rythmique sur le débit de parole à travers deux études, la première réalisée avec un amorçage rythmique musical et la seconde avec un amorçage rythmique langagier. Les changements de tempo des deux types d'amorce auditive ont provoqué un changement comparable de tempo de la vitesse d'élocution, indiquant que les amorçages rythmiques musical et langagier semblent bien s'appuyer sur des mécanismes de traitement temporel similaires, ou du moins se recouvrant entre eux.

En résumé, il ressort de ces études qu'un amorçage rythmique régulier ou dont la structure métrique est congruente avec la structure métrique de la cible langagière est bénéfique pour le traitement du langage par rapport à un amorçage irrégulier, incongruent ou une condition

³⁶ Copie vocale de l'amorce rythmique musicale avec les syllabes 'ba' et 'ka'.

³⁷ Engagement moteur (*tapping* de la main) du participant en même temps que l'amorce auditive.

sans amorçage, pour différentes populations (enfants, adultes, neurotypiques, avec un trouble développemental ou acquis du langage) et différentes langues. L'effet de l'amorçage rythmique ou métrique serait de plus renforcé par un engagement moteur du participant. L'effet comparable d'un amorçage rythmique musical ou langagier soutient l'hypothèse d'un partage de mécanismes cognitifs sous-jacents.

2.4 SYNTHÈSE : DES STRUCTURES RYTHMIQUES À LEUR TRAITEMENT NEUROPSYCHOLINGUISTIQUE EN PAROLE

Dans ce deuxième chapitre, nous avons abordé la question du rôle du rythme dans le traitement de la parole. Nous avons tout d'abord présenté les modèles neuropsycholinguistiques de traitement du langage (cf. 2.1, p. 37), en mettant en avant l'importance du couplage sensori-moteur à l'œuvre en perception comme en production de la parole. Il semblerait ainsi que les mécanismes de perception et production de la parole se co-construisent au cours du développement précoce, prédisposant ainsi à l'implication des représentations motrices de la production de parole lors de la perception, et parallèlement à l'implication des représentations auditives avec l'émission de copies d'efférence prédisant l'effet sensoriel des commandes motrices auxquelles sont comparées les sons produits en temps réel lors de la production. De plus, la proximité anatomique entre les aires motrices de contrôle des gestes impliqués dans la production de parole et les aires motrices de contrôle des gestes manuels semble indiquer que la production d'un geste manuel stimulerait l'activité motrice des gestes impliqués en production de la parole, facilitant ainsi le décodage et l'encodage de la parole.

Afin d'explorer le caractère sensorimoteur du traitement de la parole ainsi que l'implication des gestes manuels dans ce traitement, nous proposons une activité expérimentale d'amorçage rythmique de la parole (lecture de syntagmes à haute voix) mettant en comparaison différentes conditions d'amorçage manipulant ces aspects (nous développons ce questionnement sur la modalité sensorielle et le type d'amorce rythmique dans la partie suivante, p. 74).

Dans un second temps, nous avons considéré la parole en tant que signal rythmique se déroulant dans le temps à partir de l'approche de la théorie de l'attention dynamique (DAT) appliquée à la parole (Large & Jones, 1999 ; Peelle & Davis, 2012) et de l'hypothèse du rebond attentionnel (ABH, Pitt & Samuel, 1990) (cf. 2.2, p. 49). Dans cette optique, les proéminences (accents) du signal de parole structurent métriquement l'énoncé et servent de point d'ancrage aux fluctuations de l'attention de l'auditeur. Pour suivre cette pulsation, les rythmes attentionnels de l'auditeur vont se synchroniser sur les fluctuations de l'enveloppe temporelle de la parole produite par le locuteur, lui permettant de décoder plus efficacement le signal de parole entrant et de prédire l'emplacement des prochaines proéminences. Retenons également que le traitement de la pulsation engage un couplage sensori-moteur bi-directionnel entre les

régions auditives et de planification motrice : la simulation d'un mouvement engendre une activation *top-down* en direction du système auditif dans le but de prédire les battements de pulsation à venir tandis que le système auditif transmet une activation *bottom-up* en direction des aires motrices à propos de la structure métrique du stimulus externe. La perception des phénomènes rythmiques est ainsi influencée à la fois par l'attention portée à ces phénomènes qui sélectionne l'information pertinente, mais aussi par l'activité motrice de l'auditeur, qui renforce la sélection attentionnelle (Morillon et al., 2014). La perception rythmique est donc bien un processus actif. Certains auteurs préfèrent d'ailleurs parler d'expérience rythmique plutôt que de simple perception (Sauvanet 2000, Ceriani 2000, Fraise 1974).

Dans le fil de ces conceptions et modèles de traitement du rythme et du langage, l'objectif principal de notre étude consiste à caractériser l'importance de la synchronisation entre les rythmes attentionnels internes et un rythme externe, à travers deux activités expérimentales : (1) dans une première activité, nous mesurons les capacités générales de traitement rythmique et notamment les capacités de synchronisation sensorimotrice à un rythme de métronome, et (2) dans une seconde activité d'amorçage rythmique de la parole, nous mesurons l'impact spécifique de la structure métrique de l'amorce sur les capacités de traitement d'une cible langagière subséquente dont la structure métrique est ou non congruente (nous développons ce questionnement sur la congruence métrique de l'amorçage dans la partie suivante, p. 74).

Enfin, nous nous sommes intéressée au caractère supramodal du rythme (cf. 2.3, p. 61), que nous avons abordé *via* la présentation (1) de l'existence de déficits communs entre le traitement du rythme et du langage, des déficits de traitement du rythme à un âge préscolaire pouvant être précurseurs d'un trouble neurodéveloppemental du langage ; (2) de mécanismes communs entre la musique et le langage, ceux-ci pouvant mener à l'utilisation de la pratique musicale pour améliorer ou rééduquer les capacités de traitement langagières et (3) de l'impact de différentes formes d'amorçage rythmique sur la parole.

Afin d'explorer le caractère supramodal du rythme, notre étude vise plus précisément à (1) comparer les capacités générales de traitement du rythme (tâches de Tempo moteur spontané et Synchronisation sensorimotrice) avec les capacités de traitement langagières (activité d'amorçage rythmique de la parole) et (2) comparer l'effet de différentes modalités (sensorielles, sensorimotrices) et différents types (parole, musique) d'amorces rythmiques de structure métriquement congruente ou non avec la structure métrique d'une cible subséquente,

sur la production de la parole (activité d'amorçage rythmique de la parole), et ce afin de faire émerger un amorçage le plus facilitant pour la production de la parole.

II. PROBLEMATIQUE ET HYPOTHESES

En commençant ce travail de recherche, nous avons été initialement interpellée par l'existence de similitudes structurelles entre la musique et le langage et notamment l'utilisation de la musique comme outil de rééducation de troubles langagiers (cf. 2.3.2.2, p. 65). Une composante commune, le rythme, a particulièrement retenu notre attention. En effet, en tant que paramètre organisateur de la structure temporelle de la musique et de la parole, tant pour l'encodage que pour le décodage (cf. 1.2.1, p. 15), le rythme nous a semblé être un bon candidat pour explorer ces liens. Nous avons alors voulu en savoir plus sur les mécanismes sous-tendant ces liens à la croisée des domaines.

Comme évoqué en fin de partie théorique, nous nous questionnons sur le rythme en tant que faculté supramodale et transversale aux modalités sensorielles et motrices, et présentant des processus sous-jacents communs à différents domaines, tels que la musique, la danse, la marche ou encore la parole (cf. 1.1.2). En effet, le traitement de séquences rythmiques, qu'elles soient musicales, verbales ou motrices, s'effectue plus facilement à une période de référence aux alentours de 600 ms (cf. 1.2.1.3.1, p. 19). Aussi, les mouvements corporels rythmiques influencent le traitement auditif de la pulsation (cf. 2.2.2, p. 52), les capacités générales de traitement du rythme sont liées aux capacités de traitement langagières (cf. 2.3.1, p. 61) et le rythme musical, par la pratique sur le long terme ou l'amorçage (effet immédiat), influence le traitement du langage (cf. 2.3.2.2, p. 65 et 2.3.3, p. 67).

De plus, ayant choisi pour objet d'étude la parole, nous avons également voulu prendre en considération le caractère sensorimoteur du traitement langagier (cf. 2.1, p. 37). En effet, le traitement de la perception du langage s'accompagne d'une activation du réseau articulaire de production de la parole, et, inversement, la production de la parole intègre des boucles de contrôle (anticipation et rétroaction) comparant en temps réel les productions réalisées avec les productions cibles (simulation sensorielle). La motricité manuelle pourrait également participer au traitement de la parole, en raison de l'existence d'un partage anatomo-fonctionnel entre les régions motrices contrôlant les gestes manuels et les gestes de production de la parole. L'engagement moteur manuel lors du traitement langagier pourrait ainsi faciliter la perception et la production de la parole.

Ce travail se situe dans la continuité des recherches portant sur l'effet de l'amorçage rythmique sur le traitement du langage (cf. 2.3.3, p. 67), et notamment l'étude de Cason,

Astésano, et al. (2015) dont nous réutilisons ici une partie du matériel expérimental³⁸. Dans leur étude, Cason, Astésano, et al. (2015) se sont intéressés à l'effet de la congruence métrique, soit la similarité de structure métrique, entre une amorce rythmique auditive et une cible langagière, dans une tâche de détection de phonèmes. Les amorces et les cibles appartenaient à quatre groupes métriques, l'assemblage d'une amorce et d'une cible subséquente résultait ainsi en un amorçage congruent (amorce et cible appartenant au même groupe métrique) ou incongruent (amorce et cible appartenant à des groupes métriques différents) (cf. **Figure 13**).

Figure 13 :

Illustration des conditions de congruence métrique de Cason, Astésano et al. (2015)

| Condition | Amorce | Cible |
|---|--------|---|
| 1. Stress Match Number Match | | Ils dév <u>o</u> rent leur goût <u>e</u> r (x x X x x x) GM 1 |
| 2. Stress Mismatch Number Match | | Le scandale <u>e</u> x sénat (x x x X x x) GM 3 |
| 3. Stress partial Match Number Mismatch | | Il presc <u>r</u> it le bon cach <u>e</u> t (x x X x x x) GM 2 |
| 4. Stress Mismatch Number Mismatch | | Chorégr <u>a</u> phier le ballet (x x x X x x) GM 4 |

Note. Colonne de gauche : les quatre conditions métriques (1. Congruent ; 2, 3, 4, différentes granularités d'incongruence). Colonne centrale : structure métrique sous-jacente de l'amorce rythmique musicale. Colonne de droite : structure métrique sous-jacente des cibles avec un exemple de syntagme pour chaque groupe métrique possible (GM 1, 3, 2 et 4). L'intervalle entre deux « x » est de 225 ms. Les « X » en gras et majuscules représentent l'emplacement des proéminences dans la structure métrique sous-jacente. Par exemple, la condition 3. correspond à une incongruence partielle, le premier « X » occupant la même place dans la structure métrique de l'amorce que de la cible, mais les deux structures présentant un nombre différent d'éléments. GM : Groupe métrique. Adapté de Cason, Astésano et al. (2015, p. 46).

De plus, dans leur étude, les participants (adultes neurotypiques francophones) étaient répartis en deux groupes, selon la modalité de l'amorçage : pour le premier groupe, l'effet de la congruence métrique de l'amorçage était observé pour des amorces rythmiques musicales auditives (= groupe auditif) ; pour le deuxième groupe, la tâche d'amorçage rythmique musical auditif était complétée par une courte période d'entraînement audio-moteur (copie vocale du mètre de l'amorce sur les syllabes 'ba' et 'ka') (= groupe audio-moteur). Les auteurs ont trouvé qu'un amorçage congruent (condition 1. dans la **Figure 13**) résultait en des temps de réaction plus rapides (détection phonémique) que dans les conditions incongruentes (conditions 2., 3. et

³⁸ Le matériel expérimental nous a été mis à disposition par une des auteurs de l'étude et co-encadrante de cette thèse, Corine Astésano.

4. dans la **Figure 13**), cette différence n'étant significative que pour le groupe audio-moteur (tous groupes métriques confondus) et que pour les amorces des groupes métriques 1 et 2 (lignes 1. et 3. sur la **Figure 13**) (tous groupes de participants confondus).

Afin de poursuivre la réflexion engagée par Cason, Astésano, et al. (2015), nous avons souhaité d'une part nous pencher sur l'effet de l'amorçage rythmique sur le versant cette fois de la production de la parole et d'autre part questionner davantage la supramodalité du rythme et les aspects sensorimoteurs du langage dans ce processus d'amorçage rythmique. Nous avons ainsi adapté le protocole de Cason, Astésano, et al. (2015) de sorte à pouvoir comparer, en plus de l'effet de congruence métrique, l'effet de différentes modalités et différents types d'amorces sur la production de la parole. Une autre originalité de notre protocole est d'intégrer deux tâches de mesure des capacités de traitement rythmique générales (le tempo moteur spontané et la synchronisation sensorimotrice), qui seront mises en regard des observations langagières de la tâche d'amorçage rythmique (nous détaillons en détail notre protocole dans le chapitre 3, p. 86).

Plus concrètement, dans ce travail, nous avons souhaité nous pencher sur **trois questions : l'existence d'une période de référence supramodale, l'effet de l'amorçage rythmique sur la parole et les liens entre les capacités de traitement du rythme et du langage** (nous développons ces questions et leurs hypothèses correspondantes ci-après, p. 78). Ces questionnements ont été investigués auprès d'une population de 41 adultes francophones neurotypiques, qui nous permettra de mieux saisir les mécanismes sous-tendant le traitement du rythme et de la parole, ainsi que les mécanismes entrant en jeu dans le processus d'amorçage rythmique de la parole. Notre objectif est de pouvoir ensuite proposer des pistes pour la rééducation des troubles rythmiques de la parole.

Pour ce faire, nous avons élaboré un protocole expérimental visant d'une part à décrire les capacités de traitement du rythme des participants et d'autre part à décrire l'impact de l'amorçage rythmique sur la parole. Deux activités expérimentales sont ainsi proposées aux participants de l'étude.

Dans l'Activité 1, les capacités de traitement du rythme en production, perception et synchronisation sont appréhendées dans une tâche de production du Tempo Moteur Spontané (TMS) et une tâche de Synchronisation Sensori-Motrice (SMS³⁹), dans lesquelles nous observerons le TMS moyen (vitesse et stabilité) ainsi que la précision et la stabilité de la synchronisation sensorimotrice du participant. La tâche de TMS permet de déterminer si le

³⁹ De l'anglais *Sensori-Motor Synchronisation*.

participant est capable de produire une séquence rythmique et renseigne sur le tempo de préférence moteur des participants, c'est-à-dire la période de référence ; la tâche de SMS permet de déterminer si le participant est capable de percevoir une séquence rythmique à différents tempi et d'adapter son TMS en se synchronisant aux différents tempi, et renseigne sur le tempo auquel la synchronisation est la plus stable et précise. Nous détaillons l'Activité 1 en partie 3.1 (p. 86).

L'Activité 2 concerne le cœur de notre questionnement, à savoir l'effet d'un amorçage rythmique sur la production de la parole. Les participants sont ainsi invités à lire des syntagmes à haute voix (= syntagmes cibles), chaque syntagme cible à lire étant précédé d'une amorce rythmique. Dans les syntagmes produits, nous observerons les erreurs de lecture ainsi que le temps d'initiation de la parole. Les conditions expérimentales varient selon la modalité (sensorielle, sensorimotrice) et le type (vocal, musical) de l'amorce rythmique ainsi que selon la relation de congruence métrique entre l'amorce et la cible, le but étant de comparer les effets des différentes conditions et de déterminer quelle condition se trouvera être la plus facilitante pour la production de la parole. Nous détaillons l'Activité 2 en partie 3.2 (p. 89).

EXISTENCE D'UNE PERIODE DE REFERENCE SUPRAMODALE

Nous nous attendons à ce que chaque individu présente une période de référence correspondant à son tempo préféré de production, perception ou synchronisation à une pulsation rythmique (cf. 1.2.1.3.1, p. 19). Cette période de référence se situerait dans une zone allant d'environ 400 à 800 ms, avec une moyenne aux alentours de 600 ms.

Hypothèse 1 : Les mesures rythmiques recueillies dans les deux activités expérimentales sont liées entre elles et se trouvent dans une **fenêtre temporelle d'environ 400 à 800 ms et préférentiellement aux alentours de 600 ms.**

Hypothèse 1.1 : Dans l'Activité 1, le tempo auquel la Synchronisation Sensori-Motrice (SMS) est la plus stable et la plus précise est le tempo se rapprochant le plus du Tempo Moteur Spontané (TMS).

Hypothèse 1.2 : Le TMS moyen recueilli dans l'Activité 1 est corrélé au Temps d'Initiation (TI) moyen de la parole de l'Activité 2.

EFFET DE L'AMORÇAGE RYTHMIQUE SUR LA PAROLE

Nous nous attendons à ce qu'il existe un effet différentiel de l'amorçage rythmique sur la production de parole (Activité 2), selon l'engagement moteur du participant (avec ou sans *tapping*), le type d'amorce (langagier, non langagier ou musical) et la congruence métrique (structure métriquement congruente ou incongruente).

En premier lieu dans cette seconde activité expérimentale, nous souhaitons nous pencher sur l'impact de la modalité sensorielle de l'amorce rythmique. En effet, les modèles neuropsycholinguistiques et attentionnels montrent d'une part que la motricité manuelle interagit avec le traitement langagier (cf. 2.1, p. 37) et d'autre part que les traitements auditifs et moteurs interagissent lors de la perception rythmique (cf. 2.2.2, p. 52). Nous proposons ainsi de comparer l'effet d'un amorçage rythmique auditif (écoute seule) et l'effet d'un amorçage rythmique audio-moteur (écoute avec engagement moteur⁴⁰ du participant) sur la production de la parole. Nous nous plaçons ainsi dans la continuation directe des travaux de Cason, Astésano, et al. (2015) qui rapportent un effet plus important de l'amorçage rythmique audio-moteur (copie vocale de l'amorce) sur la perception de la parole par rapport à un amorçage rythmique auditif. Cason, Astésano, et al. (2015) ont ensuite formulé l'hypothèse que leurs résultats se retrouveront sur le versant de la production de la parole, du fait du couplage sensori-moteur entre perception et production de parole (cf. 2.1, p. 37). En prenant également appui sur les résultats de Falk et Dalla Bella (2016) montrant un avantage de l'amorçage rythmique audio-moteur manuel pour la perception de la parole par rapport à un amorçage auditif, **nous supposons qu'un amorçage rythmique auditif avec engagement moteur aura un effet plus facilitant sur la production de la parole qu'un amorçage rythmique auditif sans engagement moteur.**

Hypothèse 2 : Un amorçage rythmique **avec engagement moteur** des participants a un effet **positif** sur la production de parole par rapport à un amorçage rythmique auditif sans engagement moteur. Cela se manifeste par un nombre d'erreurs de lecture plus faible et un temps d'initiation de la parole plus court en condition d'amorçage rythmique avec engagement moteur.

En second lieu, nous avons voulu comparer l'effet de différents types d'amorces rythmiques sur la production de la parole. Nous avons vu dans la littérature que la production de parole peut être impactée par un amorçage langagier ou musical (cf. 2.3.3, p. 67). Or ces

⁴⁰ Mouvement du doigt (*finger tapping*).

différents amorçages n'ont jusqu'à présent pas été comparés à notre connaissance. Nous avons ainsi voulu comparer trois types d'amorces différents : une amorce de type vocal langagier, une amorce de type vocal non langagier et une amorce de type musical. Premièrement, la perception de parole activant le système moteur spécifique à la production de la parole (cf. 2.1, p. 37), **nous pouvons nous attendre à ce qu'une amorce vocale, plus incarnée, ait un effet facilitant sur la production de parole par rapport à une amorce musicale.** Deuxièmement, afin de contrebalancer la part du traitement des formes linguistiques (structures lexicales et syntagmatiques) face au traitement des formes vocales non-linguistiques dans l'amorçage, nous avons choisi d'intégrer dans cette comparaison de types d'amorce une amorce vocale non langagière (suite de syllabes simples respectant les règles phonotactiques du français). En effet, l'effet de lexicalité active automatiquement les représentations phonologiques lexicales stockées en mémoire à long terme, assurant un avantage du traitement des mots sur les pseudo-mots (Roodenrys & Stokes, 2001). Les amorces vocales langagière et non langagière activeraient ainsi toutes deux le système moteur de production de la parole, mais seule l'amorce vocale langagière activerait les représentations lexicales. De plus, la détection des prééminences dans la parole est liée à nos attentes reposant sur la syntaxe et la sémantique (Goldman et al., 2010). Le traitement d'une amorce rythmique vocale non langagière pourrait donc être moins efficace que celui d'une amorce rythmique vocale langagière, le contexte non langagier ne pouvant pas permettre à l'auditeur de formuler des attentes métriques (Shields et al., 1974). De fait, **nous supposons qu'une amorce rythmique vocale langagière sera plus facilitante pour la production de la parole qu'une amorce rythmique vocale non langagière, elle-même plus facilitante qu'une amorce rythmique musicale.**

Hypothèse 3 : L'amorçage **vocal langagier** a un effet **positif** pour la production de la parole par rapport à l'amorçage vocal non langagier, ce dernier lui-même plus facilitant que l'amorçage musical.

Hypothèse 3.1 : L'amorçage **vocal**, qu'il soit langagier ou non langagier, a un effet **positif** pour la production de parole par rapport à l'amorçage musical, car l'amorçage vocal est davantage incarné, en ce qu'il active les représentations motrices de production de la parole et donc stimule la production subséquente de parole. Cela se manifeste par un nombre d'erreurs de lecture plus faible et un temps d'initiation de la parole plus court en condition d'amorçage vocal (langagier ou non-langagier) par rapport à l'amorçage rythmique musical.

Hypothèse 3.2 : L'amorçage **vocal langagier** a un effet **positif** pour la production de parole par rapport à l'amorçage vocal non langagier : le traitement de l'amorçage langagier est automatique et stimule le traitement lexical, facilitant à la fois le traitement perceptif de l'amorce et la production lexicale subséquente, par rapport à l'amorçage vocal non langagier. Cela se manifeste par un nombre d'erreurs de lecture plus faible et un temps d'initiation de la parole plus court en condition d'amorçage vocal langagier par rapport à l'amorçage rythmique vocal non langagier.

Concernant à présent la relation de congruence métrique entre l'amorce rythmique et le syntagme cible à lire, nous nous sommes intéressée à l'effet de la similarité ou différence de structure métrique entre l'amorce et le syntagme cible, c'est-à-dire, à l'effet de congruence ou d'incongruence métrique. La perception d'une pulsation au sein d'une séquence auditive est possible grâce d'une part au mécanisme de couplage entre oscillations externes et oscillations neuronales et d'autre part à l'émergence d'une structure métrique résultant de l'imbrication et de la hiérarchisation de plusieurs périodicités dans ce couplage (cf. 2.2.1, p. 49). Ainsi, une fois les oscillations internes entraînées par une amorce, la production d'un syntagme cible subséquent devrait suivre la même structure métrique que celle perçue dans l'amorce. Or, si le syntagme cible envisagé ne présente pas la même structure métrique que celle de l'amorce perçue, le couplage oscillatoire doit être défait pour parvenir à la production métrique cible. Dans la continuation des travaux de Cason et al. (Cason, Astésano, et al., 2015 ; Cason, Hidalgo, et al., 2015 ; Cason & Schön, 2012) qui indiquent un avantage d'un amorçage rythmique métriquement congruent sur la perception et la production de parole par rapport à un amorçage métriquement incongruent, ainsi que d'après les résultats de Gould et al. (2016) ou encore Zhang et Zhang (2019) qui montrent un avantage d'un amorçage métriquement congruent sur la production de parole, **nous supposons qu'un amorçage métriquement congruent aura un effet plus facilitant sur la production de parole qu'un amorçage métriquement incongruent.**

Hypothèse 4 : L'amorçage **métriquement congruent** (structure métrique sous-jacente similaire entre le stimulus d'amorce rythmique et le syntagme cible à lire) a un effet **positif** pour la production de parole par rapport à l'amorçage incongruent. Cela se manifeste par un nombre d'erreurs de lecture plus faible et un temps d'initiation de la parole plus court en condition d'amorçage congruent par rapport à l'amorçage incongruent.

La synchronisation sensori-motrice (écoute avec *tapping*) à un stimulus rythmique auditif renforce la sélection temporelle de l'information auditive et donc l'extraction de la pulsation par rapport à une écoute sans engagement moteur (sans *tapping*) (cf. 2.2.2, p. 52). Dans la continuation des travaux de Cason, Astésano, et al. (2015) et en cohérence avec les travaux de Falk et al. (2017) qui montrent un renforcement de l'effet de congruence métrique en situation d'amorçage rythmique auditif avec engagement moteur, nous nous attendons à retrouver une **interaction** entre l'effet de la modalité sensorielle (amorce rythmique auditive avec ou sans engagement moteur) et l'effet de congruence métrique (amorçage métriquement congruent ou incongruent). **Nous supposons ainsi qu'un amorçage rythmique auditif métriquement congruent avec engagement moteur sera le plus facilitant pour la production de la parole, tandis qu'un amorçage rythmique auditif métriquement incongruent avec engagement moteur sera le moins facilitant pour la production de la parole.**

Hypothèse 5 : L'effet de la **congruence métrique** de l'amorçage est **renforcé par l'engagement moteur** des participants. Cela se manifeste par un nombre d'erreurs de lecture le plus faible et un temps d'initiation de la parole le plus court en condition d'amorçage congruent avec engagement moteur, et par un nombre d'erreurs de lecture le plus important et par un temps d'initiation de la parole le plus long en condition d'amorçage incongruent avec engagement moteur.

LIENS ENTRE LES CAPACITES DE TRAITEMENT DU RYTHME ET DU LANGAGE

Enfin, nous proposons de mettre en regard les activités expérimentales 1 (caractérisation des capacités de traitement du rythme : TMS et SMS) et 2 (amorçage rythmique de la parole). Les capacités de traitement du rythme étant considérées comme un prérequis pour le développement du langage (cf. 2.3.1, p. 61), nous proposons ainsi de comparer les données rythmiques recueillies lors de la première activité (TMS et SMS) avec les données langagières recueillies dans la seconde activité. **Nous supposons ainsi que les performances rythmiques motrices de la première activité seront liées aux performances langagières de la seconde activité.**

Hypothèse 6 : Les participants les plus stables et les plus précis lors de l'Activité 1 (TMS et SMS) font moins d'erreurs de lecture et ont un Temps d'Initiation (TI) de la parole plus court lors de l'Activité 2.

III. METHODOLOGIE EXPERIMENTALE

Cette nouvelle section est composée de cinq chapitres présentant la **méthodologie expérimentale**. Le premier chapitre (cf. 3, p. 86) concerne la présentation du plan expérimental des tâches de l'Activité 1 (Caractérisation des capacités de traitement du rythme) et de celui de l'Activité 2 (Amorçage rythmique de la parole). Nous y expliquerons les choix méthodologiques opérés ainsi que la sélection et la création du matériel expérimental.

Le deuxième chapitre (cf. 4, p. 104) concerne la procédure suivie, dans laquelle nous détaillerons les considérations et précautions éthiques respectées, le matériel (équipement technique) utilisé ainsi que le déroulé de la passation.

Le troisième chapitre (cf. 5, p. 109) nous permet de présenter la population de l'étude, les critères de sélection, la procédure de recrutement, l'estimation du nombre de participants.

Le quatrième chapitre concerne la description des données recueillies, leur nature et les pré-traitements effectués.

Enfin, le cinquième chapitre (cf. 7, p. 124) présente la démarche statistique utilisée.

3 PLAN EXPERIMENTAL

Ce chapitre est consacré à la présentation des choix méthodologiques ayant conduit à la construction des différentes tâches expérimentales des Activités 1 et 2.

3.1 PLAN EXPERIMENTAL DE L'ACTIVITE 1 : CARACTERISATION DES CAPACITES DE TRAITEMENT DU RYTHME

Les capacités de traitement du rythme sont caractérisées dans deux tâches expérimentales décrivant, l'une, la production rythmique, et l'autre, la perception et la synchronisation rythmiques.

La première tâche vise à caractériser le Tempo Moteur Spontané (TMS) (**Figure 14**), qui rend compte de la capacité du participant à produire une séquence rythmique et renseigne sur son tempo de préférence moteur. La seconde tâche concerne la caractérisation de la Synchronisation Sensori-Motrice (SMS) (**Figure 15**), qui rend compte de la capacité du participant à percevoir une séquence rythmique à différents tempi et à adapter son tempo de préférence (le TMS) en se synchronisant aux différents tempi. De plus, la SMS renseigne sur le tempo auquel la synchronisation est la plus stable et précise.

Figure 14 :
Tempo Moteur Spontané

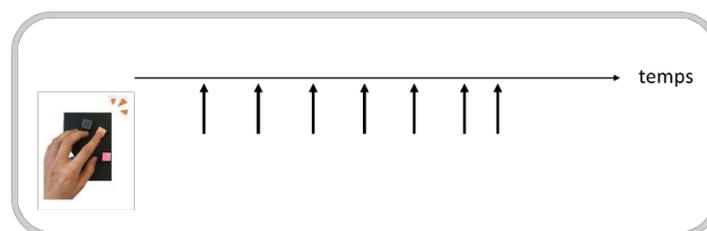
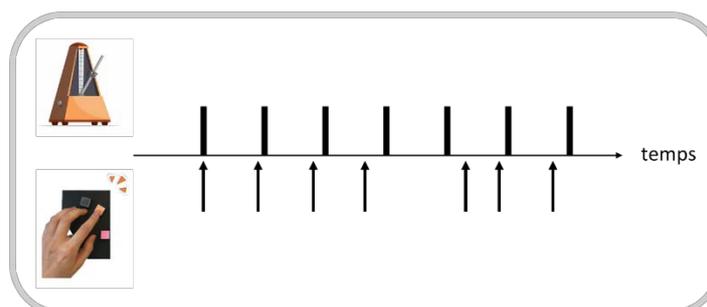


Figure 15 :
Synchronisation Sensori-Motrice



Nous détaillons ci-après le plan expérimental de la tâche de Tempo Moteur Spontané (TMS) (cf. 3.1.1) et la tâche de Synchronisation Sensori-Motrice (cf. 3.1.2, p. 87).

3.1.1 TACHE DE TEMPO MOTEUR SPONTANE

Pour rappel, la mesure du Tempo Moteur Spontané nous renseigne sur la période de référence (cf. *Hypothèse 1*, p. 78) ainsi que sur les capacités de traitement rythmiques des participants, que nous mettrons en regard des capacités de traitement langagières de l'Activité 2 (cf. *Hypothèse 6*, p. 82).

En suivant les protocoles de Tempo Moteur Spontané (TMS) de McAuley et al. (2006), Monier et Droit-Volet (2019) ou encore Trayssac (2021), préconisant de mesurer une trentaine d'Intervalles Inter-Battements (IBI⁴¹), 33 battements digitaux (*unpaced finger tapping*) ont été enregistrés, soit 32 IBI, pour la caractérisation du TMS. Cette tâche nous a permis d'obtenir deux mesures (variables dépendantes) pour chaque participant : le TMS moyen, représentatif de la période de référence du participant, et la stabilité du TMS (nous reviendrons sur le pré-traitement de ces données dans la partie 6.2, p. 114). Deux séquences de 32 IBI ont été enregistrées pour chaque participant, de sorte à obtenir deux observations pour chacune des mesures (TMS moyen et stabilité du TMS) pour chaque participant et ainsi renforcer la puissance statistique du plan expérimental.

La phase expérimentale proprement dite était précédée d'une phase d'entraînement, lors de laquelle les participants s'entraînent sur une séquence de 15 IBI (répétable à souhait).

3.1.2 TACHE DE SYNCHRONISATION SENSORI-MOTRICE

Pour la seconde partie de l'Activité 1, la caractérisation de la Synchronisation Sensori-Motrice (SMS), nous avons choisi une épreuve de battement digital (*paced finger tapping*), qui est la plus utilisée en recherche pour déterminer les capacités de SMS (pour une revue, voir Repp, 2005). Cette tâche de SMS nous renseigne sur le tempo auquel la synchronisation est la plus stable et la plus précise, que nous mettrons en relation avec la mesure du TMS moyen (cf. *Hypothèse 1*, p. 78). Les capacités de traitement moyennes des participants concernant leur stabilité et précision dans la tâche de SMS seront mises en regard des capacités de traitement langagières de l'Activité 2 (cf. *Hypothèse 6*, p. 82).

⁴¹ De l'anglais *Inter-Beat Interval*.

Le plan expérimental retenu est un plan expérimental simple, une seule variable indépendante étant observée (le tempo), avec 3 modalités expérimentales ($ISI^{42} = 225$ ms, 600 ms et 900 ms). Cette tâche nous a permis d'obtenir deux mesures (variables dépendantes) pour chaque participant : la stabilité de la SMS et la précision de la SMS (cf. 6.2, p. 114).

Selon Fraisse (1974), les intervalles de la zone de synchronisation sensori-motrice sont compris entre 200 et 1800 ms, correspondant également aux limites des durées perçues et aux limites des rythmes moteurs spontanés (cf. 1.2.1.1, p. 15). Selon Repp, cette limite se trouverait plutôt à 150 ms (Repp, 2005), voire 100 ms chez les musiciens (Repp, 2003). Nous voulions également que les tempi choisis se rapprochent des tempi utilisés pour ce type de tâche dans la littérature (par exemple, Dalla Bella et al., 2017 : 450, 600 et 700 ms ; Thompson et al., 2015 et Tierney & Kraus, 2013 : 500 et 667 ms ; McAuley et al., 2006 : 150, 225, 337, 506, 759, 1139, et 1709 ms).

Nous avons ainsi retenu trois intervalles inter-stimuli (ISI) pour les tempi de métronome : 225 ms, 600 ms et 900 ms. Le choix de ces intervalles était linguistiquement motivé, de sorte à correspondre aux ISI représentatifs de trois niveaux prosodiques (Astésano, 2022 ; Vaysse, 2023) (cf. 2.2.3.1, p. 55) :

- l'intervalle inter-syllabique moyen (225 ms)⁴³,
- l'intervalle inter-accentuel, qui correspond au niveau du mot prosodique dans notre étude (600 ms),
- l'unité intonative, qui correspond au niveau du syntagme intonatif (900 ms).

Nous avons créé plusieurs pistes de métronome afin de tester les capacités de traitement de la synchronisation rythmique audio-motrice des participants (synchronisation digitale), avec le module « Générer une piste de rythme » du logiciel Audacity (version 3.2.0) (Audacity Team, 2021). Pour chacune des pistes de métronome créées, le premier battement tombait précisément sur le début de la piste (0 s) afin de faciliter l'alignement ultérieur avec les battues des participants lors de l'analyse de la synchronisation. Les trois pistes de métronome ont été

⁴² De l'anglais *Inter-Stimulus Interval* : écart temporel entre 2 stimuli, ici 2 battements de métronome.

⁴³ 225 ms correspond par ailleurs à l'ISI moyen des amorces rythmiques utilisées pour l'Activité 2 (Amorçage rythmique de la parole).

utilisées pour la phase d'entraînement (pistes d'une durée de 15 secondes), et pour la phase expérimentale proprement dite (pistes d'une durée de 40,5 secondes⁴⁴).

Ce plan expérimental est un plan à mesures répétées, chaque participant passant par toutes les conditions expérimentales (tous les tempi). L'intérêt du plan à mesures répétées est qu'il est adapté pour peu de participants en fournissant davantage d'observations pour les différentes modalités. Il permet de plus une meilleure puissance statistique en minimisant la variabilité interindividuelle (Gillioz & Zufferey, 2020).

Lors de la phase expérimentale, chaque tempo était présenté deux fois à chaque participant afin d'augmenter le nombre d'observations pour chacune des deux variables dépendantes mesurées (stabilité et précision de la SMS). Pour limiter l'effet de transfert, l'ordre d'enchaînement des tempi était aléatorisé entre les participants, avec comme seule condition restrictive que le même tempo ne pouvait pas apparaître deux fois d'affilée.

3.2 PLAN EXPERIMENTAL DE L'ACTIVITE 2 : AMORÇAGE RYTHMIQUE DE LA PAROLE

Le questionnement au cœur de notre étude, à savoir l'impact d'un amorçage rythmique sur la facilitation de la production de parole, est mesuré dans une tâche d'amorçage de lecture à haute voix (**Figure 16**). Cette tâche se situe dans la continuation directe des travaux de Cason et al. qui ont investigué l'impact d'un amorçage rythmique sur la facilitation de la perception de la parole (détection de phonème dans des syntagmes et pseudo-mots) chez des adultes neurotypiques (Cason, Astésano, et al., 2015 ; Cason & Schön, 2012) et l'impact d'un amorçage rythmique musical sur la facilitation de la production de syntagmes (précision phonologique) chez des enfants sourds (Cason, Hidago, et al., 2015). Notre étude reprend une partie du matériel expérimental de Cason, Astésano, et al. (2015), utilisé cette fois-ci pour observer l'effet de différentes conditions d'amorçage rythmique sur la production de parole chez des adultes neurotypiques (la population de l'échantillon est décrite en partie 5, p. 109).

Cette activité nous permettra de discuter nos quatre hypothèses concernant l'effet de l'Engagement Moteur (EM) du participant, l'effet du Type d'Amorce (TA), l'effet de la Congruence Métrique (CM) de l'amorçage et l'effet d'interaction entre la congruence métrique (CM) et l'Engagement Moteur (EM) sur la facilitation de la production de la parole (cf.

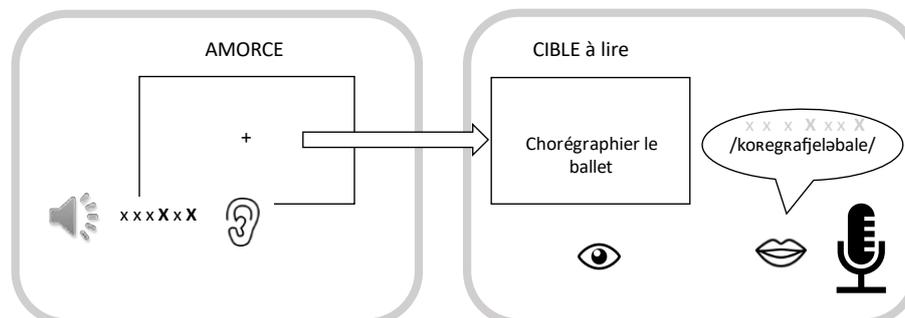
⁴⁴ Nous avons fait le choix de créer des pistes de métronome de même durée pour les 3 tempi, et nous sommes donc basée sur la durée de la piste la plus lente pour un recueil de 45 intervalles (ISI = 900 ms) (900 ms * 45 = 40,5 s).

Hypothèse 2 à 5, p. 79). La mesure du Temps d'Initiation (TI) de la parole lors de la lecture des syntagmes cibles, que nous mettrons en regard des mesures de TMS moyen et de tempo le plus stable et/ou le plus précis de la SMS (Activité 1) nous permettra de discuter l'*Hypothèse 1* (p. 78). De plus, la mesure du TI ainsi que du nombre d'erreurs de lecture seront comparées aux capacités de traitement du rythme (stabilité du TMS et stabilité et précision de la SMS Activité 1) afin de discuter l'*Hypothèse 6* (p. 82).

Dans cette partie consacrée au plan expérimental de l'Activité 2, nous présentons le paradigme expérimental de l'amorçage (cf. 3.2.1, p. 90), les variables indépendantes manipulées et variables dépendantes recueillies (cf. 3.2.2, p. 91), la sélection et la création du matériel expérimental pour les stimuli utilisés en amorces et en cibles (cf. 3.2.3, p. 95), ainsi que le processus de constitution de listes pseudo-aléatoires dans lesquelles sont répartis les participants afin d'équilibrer les différents ordres possibles de déroulé de l'activité (cf. 3.2.4, p. 102).

Figure 16 :

Amorçage rythmique de la parole (lecture à haute voix)



Note. Partie gauche : amorce rythmique auditive (le type d'amorce n'est pas spécifié) ; la séquence « x x x X x X » représente la structure métrique sous-jacente. Partie droite : cible à lire à haute voix, que nous enregistrons et recueillons. Dans cet exemple, la structure métrique sous-jacente de l'amorce (« x x x X x X ») et de la cible (« x x x X x X ») sont différentes, l'amorçage est donc métriquement incongruent.

3.2.1 LE PARADIGME DE L'AMORÇAGE

Le paradigme de l'amorçage permet de tester l'influence de la présentation d'un stimulus (« l'amorce ») sur le traitement d'un événement subséquent (« la cible »). Dans ce type de paradigme, un amorçage facilitant se traduit par des temps de réaction plus rapides et un taux d'erreurs moindre (Shao & Meyer, 2018). Notre protocole met ainsi en comparaison plusieurs amorces rythmiques (présentées dans la partie suivante 3.2.2, p. 91) afin de déterminer leur potentiel impact différentiel sur la production de syntagmes cibles. Nous avons choisi d'observer la production de parole en condition de lecture (syntagmes cibles), qui est plus contrôlée que la parole spontanée et laisse moins de place à l'expressivité et aux marques

emphatiques, permettant ainsi de nous concentrer sur l'observation de l'effet des amorces manipulées.

3.2.2 VARIABLES INDEPENDANTES MANIPULEES ET VARIABLES DEPENDANTES RECUEILLIES

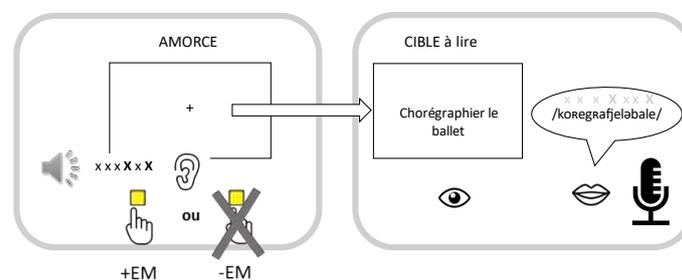
Conformément à nos *Hypothèse 2 à 5*, nous avons manipulé trois aspects (variables indépendantes – VI) de l'amorçage, à savoir (VI 1) l'Engagement Moteur (EM), (VI 2) le Type d'Amorce (TA) et (VI 3) la Congruence Métrique (CM) :

VI 1 : L'engagement moteur du participant (EM)

La manipulation de l'engagement moteur (EM) du participant (sans « -EM » ou avec « +EM ») permet de comparer l'effet d'un amorçage rythmique auditif seul (-EM) avec un amorçage audio-moteur (+EM). En situation d'amorce auditive seule (-EM), le participant a pour consigne d'écouter l'amorce de manière passive, puis de lire le syntagme cible à haute voix. En situation d'amorce audio-motrice (+EM), le participant a pour consigne d'appuyer sur une touche en même temps que chacun des sons entendus (syllabes ou sons de percussions), puis de lire le syntagme cible à haute voix (cf. **Figure 17**).

Figure 17 :

Variation de la phase d'amorce selon l'Engagement Moteur (EM) du participant



Note. La phase d'amorce (partie gauche) varie selon l'Engagement Moteur (EM) du participant : +EM (amorce audio-motrice) ; -EM (amorce auditive seule).

VI 2 : Le Type d'Amorce (TA)

En second lieu, trois types d'amorces (TA) rythmiques sont comparés (cf. **Figure 18**) :

- des amorces langagières (L), correspondant à des stimuli vocaux langagiers (par exemple, « Le scandaleux sénat ») ;
- des amorces non langagières (NL), correspondant à des stimuli vocaux non langagiers (syllabes non significantes) (par exemple, « ta ta ta ti ta ti » – les « ta »

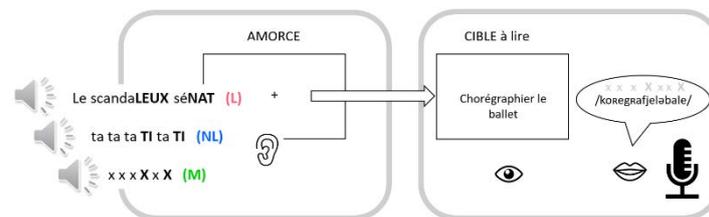
étant des syllabes métriquement faibles et les « ti » des syllabes métriquement fortes) ;

- des amorces musicales (**M**), correspondant à des stimuli musicaux (sons de percussions – les sons varient selon la force métrique des éléments dans la structure métrique).

Les différents Types d'Amorce (TA) sont détaillés dans la partie 3.2.3.2 (p. 97).

Figure 18 :

Variation de la phase d'amorce selon le Type d'Amorce



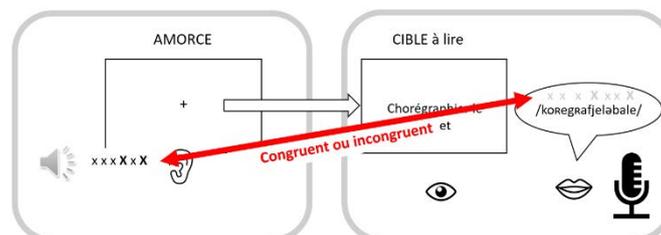
Note. La phase d'amorce (partie gauche) varie selon le Type d'Amorce (TA) : L (amorce vocale Langagière) ; NL (amorce vocale Non-Langagière) ; M (amorce vocale Musicale).

VI 3 : La congruence métrique (CM)

En dernier lieu, nous avons manipulé le lien de congruence métrique (**CM**) entre la structure de l'amorce et la structure de la cible. Lorsque la structure métrique de l'amorce et la structure métrique de la cible sont identiques, l'amorçage est congruent (**C**) ; lorsque la structure métrique de l'amorce et celle de la cible sont différentes, l'amorçage est incongruent (**IC**) (cf. **Figure 19** et **Tableau 1**). Les différentes situations de congruence sont détaillées dans la partie 3.2.3.1 (p. 95).

Figure 19 :

Variation d'amorçage selon la Congruence Métrique



Note. L'amorçage varie selon la Congruence Métrique (CM) entre la structure métrique de l'amorce (partie gauche) et la structure métrique de la cible (partie droite). Dans cet exemple, la structure métrique sous-jacente de l'amorce (« x x x X x X ») et celle de la cible (« x x x X x X ») sont différentes, l'amorçage est donc métriquement incongruent.

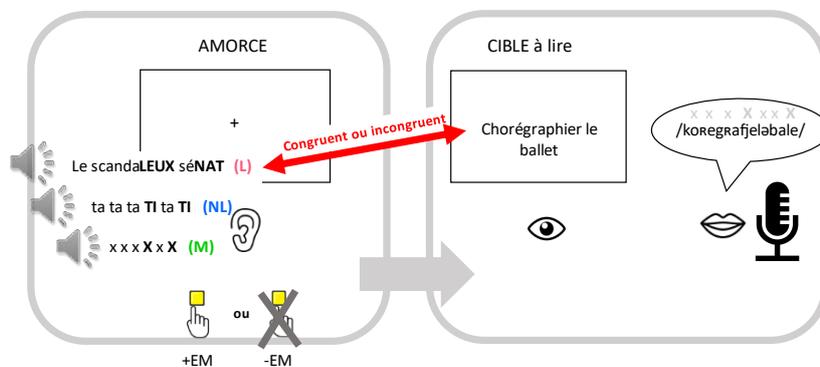
Tableau 1 :
Illustration de la congruence des couples amorce-cible

| Structure métrique de l'amorce | Structure métrique du syntagme cible | Congruence Métrique |
|--------------------------------|---|---------------------|
| x x x X x x X ⁴⁵ | x x x X x x X (par ex. : ChorégrAPHER le baLLET) | Congruent (C) |
| x x X x x X | x x x X x x X (par ex. : ChorégrAPHER le baLLET) | Incongruent (IC) |

Note. La première ligne illustre un amorçage métriquement congruent (C) (les structures métriques sous-jacentes de l'amorce et de la cible sont identiques) et la deuxième ligne un amorçage métriquement (IC) (les structures métriques sous-jacentes de l'amorce et de la cible sont différentes).

Les différentes modalités de nos trois variables indépendantes (EM, TA et CM) sont réparties en proportions égales au cours de l'expérience. Chaque participant est ainsi exposé à autant d'amorces +EM que d'amorces -EM ; autant d'amorces L que NL que M ; autant d'amorçages C que IC. De plus, les trois variables indépendantes sont cumulatives : chaque amorce est ainsi caractérisée en même temps par EM, TA et CM (par exemple, une même amorce peut être à la fois +EM, L et C) (cf. **Figure 20**). Le développement en modalités des trois variables indépendantes (qui interagissent) nous permet de construire les douze conditions du plan expérimental (cf. **Tableau 2**).

Figure 20 :
Tâche d'amorçage avec les trois Variables Indépendantes



Note. Illustration de la tâche d'amorçage, avec toutes les VI : Engagement Moteur (+EM ou -EM) ; Type d'Amorce (L, NL ou M) ; Congruence Métrique (Congruent ou Incongruent).

⁴⁵ Les « X » et syllabes en majuscules indiquent l'emplacement des syllabes/temps métriquement forts.

Tableau 2 :
Tableau récapitulatif des 12 conditions expérimentales de l'Activité 2

| Engagement Moteur du participant (EM) | Type d'Amorce (TA) | Congruence Métrique (CM) | Condition expérimentale | Nombre d'essais ⁴⁶ |
|--|-----------------------------------|-----------------------------------|-------------------------|-------------------------------|
| sans Engagement Moteur (-EM) | vocal Langagier (L) | Congruent (C) | 1 | 8 |
| | | Incongruent (IC) | 2 | 8 |
| | vocal Non Langagier (NL) | Congruent (C) | 3 | 8 |
| | | Incongruent (IC) | 4 | 8 |
| | Musical (M) | Congruent (C) | 5 | 8 |
| | | Incongruent (IC) | 6 | 8 |
| avec Engagement Moteur (+EM) | vocal Langagier (L) | Congruent (C) | 7 | 8 |
| | | Incongruent (IC) | 8 | 8 |
| | vocal Non Langagier (NL) | Congruent (C) | 9 | 8 |
| | | Incongruent (IC) | 10 | 8 |
| | Musical (M) | Congruent (C) | 11 | 8 |
| | | Incongruent (IC) | 12 | 8 |
| <i>Total pour un participant</i> | | | | 96 |

Note. L'amorçage rythmique varie selon l'Engagement Moteur du participant (EM), le Type d'Amorce (TA) et la Congruence Métrique (CM). 96 syntagmes cibles sont enregistrés par participant.

Le plan expérimental de cette activité est factoriel⁴⁷ et à mesures répétées, tous les participants passant par toutes les conditions. Afin de limiter l'impact de variables parasites comme l'ordre de présentation des conditions expérimentales et des syntagmes cibles, 72 listes d'enchaînement ont été créées⁴⁸ et administrées de manière aléatoire entre les participants (cf. 3.2.4, p. 102).

Chaque item cible n'étant lu qu'une fois par chaque participant et distribué de manière aléatoire entre les participants, le plan expérimental est également intra-item. Le contrebalancement des conditions et des items permet de limiter l'effet de transfert (effet d'ordre, effet d'apprentissage) (Gillioz & Zufferey, 2020).

Nous avons recueilli deux variables dépendantes dans cette activité : la production d'erreurs de lecture (cotée « 0 » pour un syntagme produit sans erreur de lecture ou « 1 » pour un syntagme comportant au moins une erreur) et le temps d'initiation (TI) de la parole (en ms). Nous reviendrons plus précisément sur ces variables dépendantes concernant leur nature et leur pré-traitement en partie 6.3 (p. 119).

⁴⁶ Un essai expérimental correspond à un couple amorce-cible (8 occurrences de l'essai par condition pour chaque participant).

⁴⁷ Un plan est dit factoriel lorsqu'il comporte plusieurs variables indépendantes, aussi appelées « facteurs ».

⁴⁸ Cf. Annexe 4, p. 250.

3.2.3 *MATERIEL EXPERIMENTAL UTILISE POUR LA TACHE D'AMORÇAGE RYTHMIQUE DE LA PAROLE*

Le matériel expérimental de l'Activité 2 est composé de stimuli amorces et de stimuli cibles tirés de Cason, Astésano, et al. (2015)⁴⁹. L'ensemble des syntagmes utilisés en amorces et cibles dans l'expérience se trouve en Annexe 5 (p. 244).

Dans un premier temps, nous présentons les quatre différents groupes métriques (structures métriques possibles) auxquels peuvent appartenir les amorces et les cibles (cf. 3.2.3.1), qui ont été constitués par Cason, Astésano, et al. (2015). Nous présentons ensuite les différents Types d'Amorce (TA) utilisés (cf. 3.2.3.2, p. 97) et les syntagmes cibles (cf. 3.2.3.3, p. 99). Nous terminons par l'estimation du nombre de stimuli (amorces et cibles) nécessaires pour garantir une puissance statistique suffisante (cf. 3.2.3.4, p. 100).

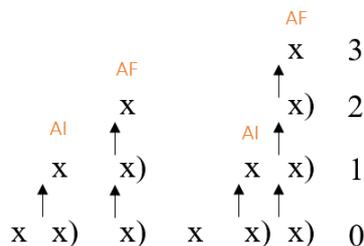
3.2.3.1 GROUPES METRIQUES

Les amorces et les cibles ont été réparties par Cason, Astésano, et al. (2015) en quatre groupes selon leur structure métrique, qui varie en fonction du nombre d'éléments constitutifs de leur structure (6 ou 7) et de la place du premier accent final (ou temps métriquement fort) (en troisième ou quatrième position dans la structure). Plus précisément, chaque stimuli est ainsi constitué de deux syntagmes accentuels (qui peuvent souvent correspondre au niveau des mots prosodiques). Le premier est composé de 3 ou 4 éléments/syllabes, correspondant à la taille moyenne du mot prosodique en français (Astésano, 2001, p. 34) et le second de 2, 3 ou 4 éléments/syllabes (cf. **Figure 21** et **Tableau 3**).

⁴⁹ Le matériel expérimental nous a été mis à disposition par une des auteurs de l'étude et co-encadrante de cette thèse, Corine Astésano.

Figure 21 :

Représentation en grille métrique de l'énoncé « Ils dévorent leur goûter »



Ils dévorent leur goûter

Note. Le syntagme « Ils dévorent leur goûter » est extrait du matériel expérimental de Cason, Astésano et al. (2015), que nous réutilisons dans le cadre de ce travail de thèse pour les stimuli amorces et cibles. Le premier AF a un poids métrique moins important que le second AF, ce dernier étant projeté plus haut dans la hiérarchie métrique. Les « x) » renvoient aux syllabes proéminentes. 0 : interface entre le plan segmental et le plan métrique (syllabes) ; 1 : pied ; 2 : syntagme accentuel ; 3 : syntagme intonatif. AI : Accent Initial ; AF : Accent Final. Adapté de Di Cristo (2016, p. 61).

Tableau 3 :

Illustration des quatre patrons de groupes métriques pour les amorces L et NL

| | Groupe métrique 1 | Groupe métrique 2 | Groupe métrique 3 | Groupe métrique 4 |
|----|---------------------|--------------------------|---------------------------|------------------------|
| | x x x X x X (n = 6) | x x X x x X (n = 6) | x x X x x x X (n = 7) | x x x X x x X (n = 7) |
| L | Le scandaLEUX sénAT | Ils déVORENT leur goûTER | Il presCRIT le bon caCHET | ChorégrAPHER le baLLET |
| NL | ta ta ta TI ta TI | ta ta TI ta ta TI | ta ta TI ta ta ta TI | ta ta ta TI ta ta TI |

Note. Les « X » et syllabes en majuscules indiquent l'emplacement des syllabes métriquement fortes.

Ces quatre groupes métriques permettent de faire varier les modalités expérimentales de la Congruence Métrique (CM) des couples 'amorce-cible' (pour une illustration de CM, se rapporter au **Tableau 1**, p. 93). A la différence de Cason, Astésano, et al. (2015) dont le plan expérimental comportait quatre conditions de congruence⁵⁰ (cf. **Figure 13**, p. 76), nous n'avons retenu que deux conditions de congruence : congruent ou incongruent (correspondant aux conditions 1 et 4 dans l'étude de Cason, Astésano, et al., 2015). En effet, notre objectif de recherche étant principalement porté sur l'effet de l'Engagement Moteur (EM) et l'effet différentiel des Types d'Amorce (TA), nous ne pouvions pas conserver la même granularité concernant la Congruence Métrique que dans l'étude de Cason, Astésano, et al. (2015), pour des raisons de faisabilité du plan expérimental qui aurait sinon été surchargé. Cependant, afin

⁵⁰ Conditions 1. Congruent (structures métriques identiques entre l'amorce et la cible) ; 2. Incongruence partielle (même nombre d'éléments entre la structure de l'amorce et de la cible mais déplacement du premier accent final) ; 3. Incongruence partielle (même emplacement du premier accent final entre la structure de l'amorce et de la cible mais nombre d'éléments différent) ; 4. Incongruence complète (nombre d'éléments et emplacement du premier accent final différents entre la structure de l'amorce et de la cible).

de pouvoir avoir un nombre suffisant de stimuli amorces et cibles (cf. 3.2.3.4, p. 100), nous avons quand même pris en considération le groupe métrique de sorte à contrôler cette variable.

Ainsi, pour constituer un couple métriquement Congruent (C), une amorce du groupe métrique 1 devra être appariée avec un syntagme cible du groupe métrique 1. Pour constituer un couple InCongruent (IC), une amorce du groupe métrique 1 devra être appariée avec une cible du groupe métrique 3, différant ainsi à la fois en nombre d'éléments et en place de leur premier accent final/temps fort dans leur structure.

3.2.3.2 AMORCES RYTHMIQUES

Dans cette tâche d'amorçage rythmique, nous avons souhaité comparer l'impact de différents types d'amorces sur la parole. Bien qu'il existe plusieurs études rapportant que la production de parole peut être impactée par différents types d'amorçages rythmiques auditifs – amorçage de parole (voir par exemple Aichert et al., 2021 ; Jungers et al., 2016), amorçage musical (voir par exemple Cason, Hidalgo, et al., 2015 ; Gould et al., 2016 ; Jungers et al., 2016) –, l'impact différentiel de ces amorçages n'est à notre connaissance pas comparé. C'est donc ce que nous avons souhaité mettre en place dans ce travail de recherche en comparant l'impact d'amorces langagières et musicales sur la production de parole. Nous avons également souhaité intégrer dans cette comparaison la part de l'effet de lexicalité de l'amorce, avec la création d'un amorçage vocal non langagier.

Les stimuli utilisés en tant qu'amorces rythmiques auditives sont donc de trois types : vocal Langagier (L), vocal Non Langagier (NL) et Musical (M). Les amorces L et M sont tirées du matériel expérimental de Cason et al. (2015), tandis que les amorces NL ont été créées pour la présente expérience. Toutes les amorces (L, NL et M) sont équitablement réparties dans les quatre groupes métriques (cf. 3.2.3.1, p. 95). L'intervalle inter-stimuli (ISI) moyen pour chacun des trois types d'amorce est de 225 ms (4,44 Hz), équivalent d'un intervalle inter-syllabique moyen en parole naturelle (cf. **Figure 13**, p. 76).

Les amorces musicales (M) sont des sons de percussions créés dans Adobe Audition dans le cadre de l'étude de Cason, Astésano, et al. (2015). Les sons de percussion inaccentués ont un timbre de *hi-hat* (charleston) fermé, une durée de 138 ms et une valeur de *Root Mean Square* (RMS) de -43,63 dB. Le son correspondant à l'accent final du premier pied métrique a un timbre rimshot, une durée de 196 ms et une valeur de RMS de -41,25 dB. Le son correspondant à l'accent final du second pied métrique a un timbre *snare*, une durée de 353 ms

et une valeur de RMS de -28,88 dB. La variation de durée entre les sons de percussions permet de respecter la hiérarchie métrique (cf. **Figure 21**, p. 96), les sons inaccentués étant plus courts que le premier accent final, celui-ci plus court que le second accent final. De plus, les durées des sons de percussions inaccentués et accentués renvoient aux durées moyennes de syllabe inaccentuée, accent final de syntagme et accent final d'unité intonative (Astésano, 2001 ; Delattre, 1966 ; Fant et al., 1991).

Les amorces vocales langagières (L) (par exemple : « Le scandaleux sénat ») ont été lues à haute voix par une locutrice francophone native (C. Astésano, co-auteurice de l'étude de Cason, Astésano, et al., 2015) dans une salle insonorisée et ont été créées dans le cadre de l'étude de Cason, Astésano, et al. (2015). Les syntagmes étaient prononcés au sein d'une phrase porteuse dont seule la première partie était conservée. Le syntagme conservé était prononcé avec une modalité de continuation majeure à la fin. Les mots des syntagmes ont été sélectionnés afin de contrôler leur fréquence lexicale. La structure métrique réalisée était calquée temporellement et métriquement sur la structure des amorces musicales (M), de sorte à conserver en moyenne les mêmes durées d'intervalles et de syllabes (inaccentuées, AF de syntagme, AF d'unité intonative).

Les amorces vocales non langagières (NL) sont des amorces vocales créées pour la présente étude et prononcées par la même locutrice francophone que les amorces L de l'étude de Cason, Astésano, et al. (2015). Les stimuli NL correspondent à des syllabes simples (consonne + voyelle) respectant les règles phonotactiques du français. Les syllabes inaccentuées et accentuées se différencient par leur timbre vocalique (« ta » pour les syllabes inaccentuées et « ti » pour les syllabes accentuées). Les stimuli NL (par exemple : « ta ta ta TI ta TI ») ont été enregistrés dans une salle insonorisée avec un enregistreur TASCAM (DR-40). Tout comme pour les amorces langagières (L), la structure métrique réalisée pour les amorces vocales non langagières (NL) était calquée temporellement et métriquement sur la structure des amorces musicales (M), de sorte à conserver en moyenne les mêmes durées d'intervalles et de syllabes (inaccentuées, AF de syntagme, AF d'unité intonative).

Afin de renforcer l'effet d'entraînement de l'amorce sur la cible, les amorces (L, NL et M) ont été manipulées pour la présente étude avec le logiciel Audacity de manière à comporter six répétitions de chaque stimulus (cf. **Figure 22**). En effet, selon Fraisse (1966), au moins trois répétitions d'un stimulus ou d'une structure simple sont nécessaires pour établir une synchronisation. Pour Pitt et Samuel (1990), les attentes rythmiques seraient plutôt pleinement

établies à partir de la cinquième occurrence d'un battement (ou d'une syllabe accentuée). Ce nombre de six répétitions de l'amorce que nous retenons est par ailleurs cohérent avec d'autres protocoles de recherche testant l'effet d'un amorçage rythmique sur le traitement du langage (par exemple, Bedoin et al., 2016 : 8 répétitions de l'amorce ; Canette et al., 2019 : 4 répétitions ; Cason & Schön, 2012 : 4 répétitions ; Gould et al., 2016 : 2 répétitions ; Zhang & Zhang, 2019 : 3 répétitions).

Figure 22 :

Visualisation d'une amorce de type musical dans le logiciel Audacity



Note. Visualisation d'une amorce de type musical (M) répliquée six fois dans le logiciel Audacity. L'amorce présente une structure métrique de 6 éléments, avec le premier temps fort tombant en quatrième position.

Nous avons donc inséré une pause silencieuse entre deux amorces répétées de sorte à délimiter le patron métrique répété, la durée de cette pause étant proportionnelle à la durée des intervalles inter-syllabiques moyens (225 ms) afin de renforcer l'effet d'entraînement. Ainsi, pour les amorces comportant 6 syllabes/sons, nous avons ajouté une pause silencieuse afin que la durée entre le dernier AF d'unité intonative et le premier son du stimulus suivant soit de 675 ms (3×225 ms) ; pour les amorces comportant 7 syllabes/sons, nous avons ajouté une pause silencieuse afin que la durée entre le dernier AF d'unité intonative et le premier son du stimulus suivant soit de 900 ms (4×225 ms). Les amorces résultantes (avec six répétitions des amorces initiales) sont donc respectivement de 10.800 ms (pour les amorces ayant initialement 6 syllabes/sons) et 13.500 ms (pour les amorces ayant initialement 7 syllabes/sons).

3.2.3.3 SYNTAGMES CIBLES

Les syntagmes cibles (syntagmes à lire à voix haute) sont également issus du matériel expérimental de Cason et al. (2015), qui les ont construits suivant le même principe que les stimuli d'amorce, soit des syntagmes composés de 2 syntagmes accentuels (6 ou 7 syllabes au total).

3.2.3.4 ESTIMATION DU NOMBRE DE STIMULI

L'estimation du nombre de stimuli a été conduite par diverses contraintes : des contraintes statistiques, imposant un nombre minimal de stimuli, et des contraintes propres au plan expérimental.

En tenant compte des contraintes de puissance statistique liées à la taille de notre échantillon (cf. 5.3, p. 111), nous avons une exigence de recueillir 8 essais par condition expérimentale (pour un rappel des 12 conditions expérimentales, se rapporter au **Tableau 2**, p. 94). Il nous fallait ainsi 96⁵¹ couples amorce-cible, soit 96 amorces (32 amorces L, 32 amorces NL et 32 amorces M) et 96 syntagmes cibles par passation expérimentale pour un participant.

Une autre contrainte importante du plan expérimental concernait les stimuli langagiers : un même syntagme (amorce L ou syntagme cible), par exemple « Le scandaleux sénat », ne pouvait apparaître qu'une seule fois dans l'expérience pour éviter des effets d'apprentissage, et donc ne pouvait être utilisé qu'en amorce ou qu'en cible. Il nous fallait ainsi un total de 32 syntagmes différents en amorces L + 96 syntagmes cibles différents, soit 128 syntagmes différents pour la phase expérimentale proprement dite.

Une dernière contrainte concernait la présence d'une phase d'entraînement, préalable à la phase expérimentale proprement dite. Pour faciliter le déroulé de la passation (sur lequel nous reviendrons en partie 4.3.2, p. 107), l'Activité 2 a été divisée en six blocs expérimentaux (cf. **Figure 23**, p. 101) : deux blocs principaux selon l'Engagement Moteur (EM) du participant, et trois sous-blocs à l'intérieur des deux blocs principaux selon le Type d'Amorce (TA) (nous détaillons les règles d'enchaînement des blocs dans la partie suivante, 3.2.4, p. 102). Chacun des six blocs démarrait ainsi par une courte phase d'entraînement (1 essai par condition expérimentale par sous-bloc), suivie de la phase expérimentale proprement dite. Les syntagmes utilisés dans les phases d'entraînement devaient ainsi être différents des syntagmes de la phase expérimentale proprement dite. Pour la phase d'entraînement, nous avons ainsi besoin de 2 amorces L et 6 syntagmes cibles dans le bloc -EM, et de 2 amorces L et 6 syntagmes cibles dans le bloc +EM, soit un total de 16 syntagmes.

Nous avons ainsi besoin au total pour une passation (un participant) d'une base de 144 syntagmes⁵², répartis entre les syntagmes utilisés en amorces (pour les amorces de type

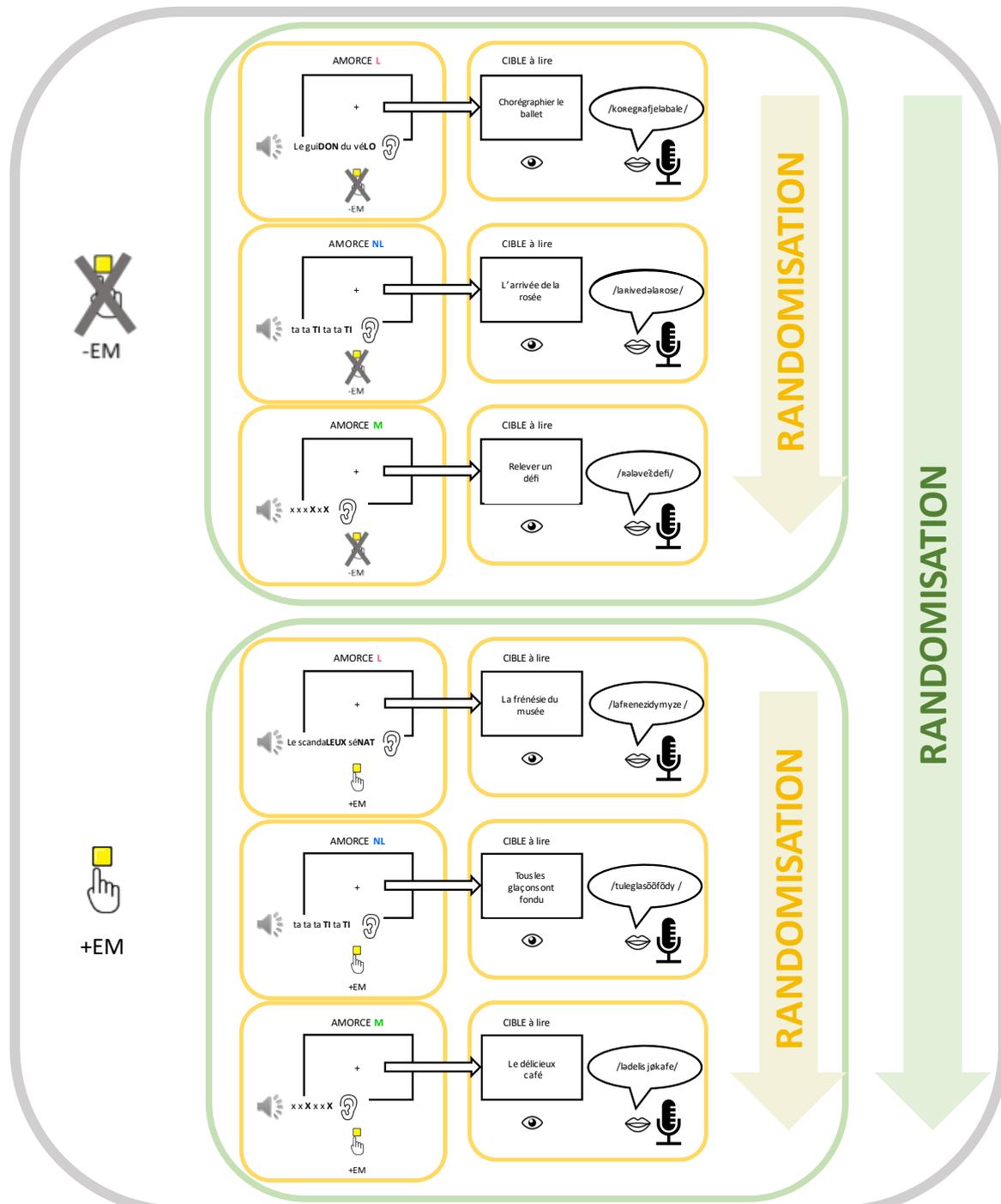
⁵¹ 12 conditions * 8 essais

⁵² 16 syntagmes pour la phase d'entraînement + 128 syntagmes pour la phase expérimentale proprement dite = 144 syntagmes.

Langagier) et les syntagmes cibles à lire par le participant, et ce pour la phase d'entraînement et la phase expérimentale proprement dite. L'ordre des syntagmes était aléatorisé entre les participants.

Figure 23 :

Illustration des six blocs expérimentaux de l'Activité 2



Note. Illustration des six blocs expérimentaux : deux blocs principaux selon l'Engagement Moteur du participant (-EM et +EM), dont l'ordre est randomisé selon les participants ; trois sous-blocs selon le Type d'Amorce (Langagier – L ; Non Langagier – NL ; Musical – M), dont l'ordre est randomisé (à l'intérieur d'un bloc -EM ou +EM) selon les participants.

3.2.4 CONSTITUTION DES LISTES PSEUDO-ALEATOIRES

Afin de renforcer la validité interne du plan expérimental, nous avons opté pour différents niveaux d'aléatorisation de nos variables indépendantes et items expérimentaux : aléatorisation de l'ordre des blocs expérimentaux et aléatorisation des amorces et des cibles, ce que nous avons contrôlé par la création de listes pseudo-aléatoires.

Les listes pseudo-aléatoires ont été constituées en respectant une répartition des couples amorce-cible en différents blocs liés aux conditions expérimentales (cf. **Figure 23**). Le fait de découper l'expérience en plusieurs blocs a permis de simplifier les consignes et le déroulement de l'expérience. Deux principaux blocs ont été constitués selon la condition « Engagement Moteur du participant » (un bloc -EM et un bloc +EM), dont l'ordre d'enchaînement a été aléatorisé entre les participants. Ces deux blocs principaux étaient constitués chacun de trois sous-blocs, suivant le Type d'Amorce (sous-blocs L, NL et M), dont l'ordre d'enchaînement a à nouveau été aléatorisé entre les participants. Chaque passation a ainsi été divisée en six blocs expérimentaux, à l'intérieur desquels la condition expérimentale Congruence Métrique et les stimuli utilisés en amorces et items cibles ont été aléatorisés entre les participants.

Les six blocs possibles sont les suivants : +EM_L, +EM_NL, +EM_M, -EM_L, -EM_NL, -EM_M. Afin de faciliter le déroulé de la passation et l'appropriation des consignes pour les participants, la contrainte suivante a été posée : les trois sous-blocs (type d'amorce L, NL et M) renvoyant à la même modalité expérimentale d'engagement moteur (+EM ou -EM) ne peuvent pas être séparés. Ainsi par exemple, l'enchaînement suivant pour le déroulé d'une passation expérimentale n'est pas possible puisque les modalités d'engagement moteur se mélangent : +EM_L, -EM_M, +EM_NL, -EM_L, +EM_M, -EM_NL. Connaissant le nombre de blocs (3*2) et la contrainte d'enchaînement des blocs relative à l'engagement moteur, le nombre de combinaisons possibles d'enchaînement des six blocs est égal à :

$$3! \cdot 3! \cdot 2 = 72$$

Afin de limiter les effets qui auraient pu être liés à l'ordre des blocs, nous devons donc piocher dans ces 72 possibilités de combinaisons pour que tous les participants soient exposés à un enchaînement différent des parties expérimentales (blocs).

En second lieu, une fois les 72 listes d'enchaînements possibles établies, il fallait les remplir avec les stimuli d'amorces et les syntagmes cibles, tout en respectant certaines contraintes pour limiter les effets liés à l'ordre des items utilisés comme amorces ou comme cibles. Conformément aux contraintes statistiques liées à notre plan expérimental (cf. 5.3, p.

111), nous avons sélectionné 8 essais expérimentaux par condition expérimentale (cf. **Tableau 2**, p. 94), soit 96 (12*8) essais expérimentaux par participant pour la phase expérimentale proprement dite. Nous avons également sélectionné 1 essai d'entraînement par condition expérimentale, soit 12 essais d'entraînement pour l'ensemble des phases d'entraînement. Il fallait ainsi répartir au total 108 essais entre les phases d'entraînement et les phases expérimentales tout au long d'une passation. Il a également fallu contrôler la congruence métrique des couples amorces-cibles : les amorces et les cibles ne pouvaient pas simplement être distribuées de manière aléatoire, mais devaient respecter une répartition équilibrée entre couples amorces-cibles métriquement congruents et incongruents. De même, nous avons contrôlé le fait que les quatre groupes métriques (cf. 3.2.3.1, p. 95) soient répartis de manière équilibrée entre les 12 conditions expérimentales. Enfin, nous avons aussi posé comme règle d'enchaînement que trois amorces du même groupe métrique, trois cibles du même groupe métrique, trois essais congruents ou trois essais incongruents ne pouvaient pas apparaître d'affilée pour éviter un effet d'apprentissage.

Parmi les 72 listes qui ont été créées, les 36 premières listes correspondaient à un ordre expérimental « -EM_+EM »⁵³ et les 36 dernières listes correspondaient à un ordre expérimental « +EM_-EM ». Nous prenions donc soin d'alterner les listes « -EM_+EM » et « +EM_-EM » entre les participants au fur et à mesure pour que les données recueillies soient équilibrées et ne subissent pas d'effet d'ordre de présentation des parties de l'expérience.

Les 72 listes ont été créées sur le logiciel Excel⁵⁴ et ont été implémentées dans notre application PsychoPy pour ordonner le déroulé de chaque passation (cf. Annexe 4, p. 243).

⁵³ Les participants piochant une de ces 36 premières listes « -EM_+EM » commençaient donc la passation par les trois blocs expérimentaux (L, NL et M – ordre aléatoire) dans la condition « sans Engagement Moteur » (-EM) puis enchaînaient avec les trois blocs expérimentaux (L, NL, M – ordre aléatoire) dans la condition « avec Engagement Moteur » (+EM).

⁵⁴ Nous remercions C. Gindre (société toulousaine CGExcel) pour son aide conséquente dans la réalisation des listes d'aléatorisation.

4 PROCEDURE

Dans ce chapitre, nous présentons les considérations et précautions éthiques respectées (cf. 4.1), le matériel technique utilisé (cf. 4.2) ainsi que le déroulé de la passation (cf. 4.3, p. 106).

4.1 CONSIDERATIONS ETHIQUES ET ANONYMISATION DES DONNEES

L'étude présentée dans ce travail de recherche a été approuvée par le Comité d'Ethique de la Recherche (CER) de Toulouse (avis n°2022_545) ainsi que par le délégué à la protection des données de l'Université Toulouse 2 (avis n°R-202205161732).

Le consentement éclairé des participants a été recueilli avant de démarrer l'étude, après que les consignes leur aient été exposées (cf. Annexe 1, p. 236).

Toutes les données sont rendues anonymes par l'affectation d'un identifiant aléatoire à chaque participant. Afin de garantir l'anonymat des participants, dans le respect du Règlement Général sur la Protection des Données (RGPD), chaque participant a reçu un identifiant unique à son arrivée. Cet identifiant était conçu de la manière suivante : agglomération d'une lettre et d'un numéro à deux chiffres. Le numéro à deux chiffres était donné par le numéro des listes pseudo-aléatoires constituées pour l'Activité 2 (cf. Annexe 4, p. 243) (liste 01 à liste 72), auquel a été rajoutée une lettre de l'alphabet (participant P01, Q02, R03 et ainsi de suite jusqu'à I72). Cet identifiant était ensuite lié aux données du participant recueillies lors de l'Activité 1 et de l'Activité 2, ainsi que dans un questionnaire socio-démographique (informations générales, cf. 6.1, p. 113). Une étiquette portant l'identifiant était collée sur l'exemplaire de la notice d'information et du formulaire de consentement que le participant rapportait chez lui après l'expérience. Ainsi, après la passation, seul le participant était en possibilité de faire le lien entre son identité et son identifiant. Aucune table de correspondance n'a été conservée pour les besoins de l'étude.

4.2 MATERIEL TECHNIQUE

Chaque passation expérimentale individuelle s'est déroulée au plateau technique PETRA⁵⁵ de la plateforme Cognition, Comportements et Usages de l'Université Toulouse 2.

⁵⁵ Plateau d'Etudes Techniques et de Recherches en Audition.

Ce plateau nous a permis d'avoir à disposition une pièce insonorisée et du matériel technique permettant la diffusion et la captation sonore (cf. **Figure 24**).

Les activités 1 et 2 se sont déroulées sur un ordinateur iMac (MacOS Mojave 10.14.6). Les stimuli sonores ont été diffusés via trois haut-parleurs disposés devant et sur les côtés du participant. Un boîtier de réponse⁵⁶ (carte Arduino Pro Micro à 1000Hz de fréquence d'échantillonnage sur support imprimé en 3D) à 6 touches mécaniques a servi de clavier afin d'enregistrer les frappes digitales (*tapping*). La parole a été enregistrée *via* un microphone-casque (C520 Vocal Condenser Microphone, AKG Acoustics). Un réglage des haut-parleurs et du niveau d'entrée du microphone a été effectué en amont de l'expérience pour chaque participant afin d'assurer une qualité d'écoute puis de traitement des données optimale.

Figure 24 :
Salle de passation expérimentale



Les activités expérimentales ont été construites avec le logiciel PsychoPy (v2022.2.4) (Peirce et al., 2019). Après avoir comparé différentes interfaces expérimentales (PERCEVAL, LANCELOT, JsPsych, E-Prime, OpenSesame et PsychoPy) sur lesquelles des versions bêta de protocole ont été construites et pré-testées, nous avons finalement retenu PsychoPy pour ses nombreux avantages. En premier lieu, PsychoPy permet d'intégrer tous les éléments nécessaires pour nos deux activités expérimentales, soit notamment la possibilité de jouer des sons, afficher du texte, recueillir et exporter des données comportementales (enregistrements sonores, boutons du boîtier utilisés, temps de réaction), contrebalancer l'ordre des parties de l'expérience en tenant compte de nos listes pseudo-aléatoires (cf. Annexe 4, p. 243). De plus, PsychoPy fournit une grande précision temporelle, nécessaire pour les mesures comportementales (Bridges et al., 2020). Nous avons ensuite été attentive au côté pratique du logiciel : celui-ci peut être installé entre autres sur Mac (condition imposée par le matériel de la salle de passation)

⁵⁶ Prêté par Dimitri Paisios, doctorant au laboratoire CLLE (Toulouse 2).

et peut fonctionner avec des dispositifs externes (boîtier de réponse externe par exemple dans notre cas). La construction de l'application expérimentale est possible via une interface « Builder »⁵⁷, qui facilite le montage de l'expérience à l'aide de modules pré-codés (par exemple, module « microphone », module « texte » etc.) et est tout de même assez souple de manière à permettre aussi l'insertion de code (Python) lorsque les modules proposés ne sont pas suffisants. PsychoPy a aussi l'avantage d'être un logiciel open source avec des mises à jour régulières, une documentation riche et un important forum d'entraide où échangent les utilisateurs et les concepteurs (<https://discourse.psychopy.org/>).

4.3 DEROULE

A son arrivée sur le lieu de passation, l'adéquation du participant aux critères de sélection est vérifiée. Il lui est remis une fiche d'information sur le déroulé de la passation et les consignes ainsi qu'un formulaire de consentement en double exemplaire à signer pour prendre part à l'étude (cf. Annexe 1, p. 236). Chaque participant reçoit un identifiant aléatoire qui garantit son anonymat.

Le participant prend part à deux activités expérimentales, pour une durée totale d'environ 45 minutes :

- Activité 1 : Caractérisation des capacités de traitement du rythme (nous détaillons le déroulé de l'Activité 1 dans la partie suivante 4.3.1, p. 107)
- Activité 2 : Amorçage rythmique de la parole (nous détaillons le déroulé de l'Activité 2 dans la partie 4.3.2, p. 107)

Les différentes consignes sont une première fois expliquées à l'oral, puis sont rappelées à l'écrit au fil de l'expérience.

Le participant complète un questionnaire sociodémographique entre les deux activités expérimentales (cf. Annexe 2, p. 239).

En remerciement pour sa participation, chaque participant a reçu une place de cinéma.

⁵⁷ Il existe également une interface « Coder » pour les codeurs expérimentés.

4.3.1 *DEROULE DE L'ACTIVITE 1 : CARACTERISATION DES CAPACITES DE TRAITEMENT DU RYTHME*

L'activité 1 comprend deux tâches : la tâche de Tempo Moteur Spontané (TMS) et la tâche de Synchronisation Sensori-Motrice (SMS), dont nous détaillons le déroulé ci-après.

4.3.1.1 CARACTERISATION DU TEMPO MOTEUR SPONTANE

Pour la caractérisation du Tempo Moteur Spontané (TMS), il est demandé aux participants de battre la pulsation à une vitesse qui leur semble la plus naturelle et la plus agréable (Fraisie, 1974), en appuyant sur une touche du boîtier de mesure. L'expérience démarrait avec une phase d'entraînement pour se familiariser avec la consigne, répétable à souhait, avant de commencer la phase expérimentale proprement dite, où le recueil du TMS se faisait sur deux séquences séparées de quelques secondes de pause.

4.3.1.2 CARACTERISATION DE LA SYNCHRONISATION SENSORI-MOTRICE

La consigne donnée aux participants était de se synchroniser aux différents tempi de métronome (225, 600 et 900 ms) en appuyant sur une touche en même temps que chaque son de métronome entendu. L'expérience démarrait avec une phase d'entraînement, répétable à souhait, avant de commencer la phase expérimentale proprement dite, où chaque tempo est présenté deux fois au participant dans un ordre aléatoire.

4.3.2 *DEROULE DE L'ACTIVITE 2 : AMORÇAGE RYTHMIQUE DE LA PAROLE*

Dans cette tâche, les participants ont pour consigne de lire tout haut des syntagmes courts présentés sur un écran d'ordinateur, et ce dès l'apparition du syntagme. Chaque syntagme à lire (cible) est précédé d'un stimulus rythmique (amorçage) auditif (-EM) ou audio-moteur (+EM), langagier (L), non langagier -NL) ou musical (M), métriquement congruent (C) ou incongruent (IC). Les syntagmes lus par les participants sont enregistrés.

Afin de prévenir la fatigabilité des participants, cette seconde activité expérimentale est divisée en plusieurs blocs (cf. **Figure 23**, p. 101) entrecoupés de pauses à la discrétion du participant. Chacun des blocs démarre par un court entraînement suivi de la phase expérimentale proprement dite, pour que les participants se familiarisent avec la consigne. Lors de la phase d'entraînement, le participant s'entraîne sur un essai métriquement congruent et un essai incongruent. Il peut ensuite choisir de répéter la phase d'entraînement autant de fois qu'il le souhaite. Lors de la phase expérimentale proprement dite, chaque condition expérimentale apparaît huit fois (= huit essais expérimentaux), de sorte à avoir un nombre d'observations

suffisant pour rendre compte de l'effet de chaque condition expérimentale (cf. **Tableau 2**, p. 94).

5 POPULATION

Dans ce chapitre, nous présentons les critères de sélection (inclusion et exclusion) des participants (cf. 5.1), le mode de recrutement (cf. 5.2, p. 110), comment a été estimé le nombre de participants (cf. 5.3, p. 111) et terminons par la description de l'échantillon (cf. 5.4, p. 111).

5.1 CRITERES DE SELECTION DES PARTICIPANTS

Afin de mieux saisir les mécanismes généraux entrant en jeu dans le processus d'amorçage rythmique de la parole, l'étude a été conduite auprès d'une population neurotypique.

Pour que la population de l'échantillon soit la plus homogène possible, nous avons posé certaines limites de sélection. Concernant l'âge des participants, nous avons ciblé la tranche adulte puisque la variabilité interindividuelle peut être plus importante dans les autres tranches d'âge pour certaines tâches expérimentales envisagées. Concernant la tâche de Synchronisation Sensori-Motrice de l'Activité 1 par exemple, la variabilité de la synchronisation décroît au fil de l'enfance et de l'adolescence pour se stabiliser à l'âge adulte, avant d'augmenter de nouveau chez les personnes âgées (McAuley et al., 2006 ; Repp, 2005 ; Thompson et al., 2015). Concernant l'Activité 2 de lecture après amorçage rythmique, l'âge a également un effet sur la vitesse de lecture, qui augmente lors de l'enfance jusqu'à l'âge adulte, avant de diminuer à nouveau au-delà de 40 ans (A. H. Chen et al., 2019).

Nous avons choisi de ne pas inclure les professionnels de la musique ou de la danse, en raison de leur expertise rythmique (Provasi, 2016), qui pourrait se refléter dans leurs performances (cf. 2.3.2). En effet, les musiciens et les danseurs auraient ainsi plus de facilité à synchroniser leurs mouvements corporels à un rythme que les non-musiciens ou non-danseurs (pour la musique, voir Drake, Jones, et al., 2000 ; Repp, 2005 ; Thompson et al., 2015 ; pour la danse, voir Miura et al., 2011), bien que ces résultats différentiels de synchronisation liés à l'expertise ne soient pas retrouvés chez tous les auteurs (voir par exemple Large et al., 2002). De plus, la pratique musicale aurait un effet bénéfique sur le traitement du langage (Patel, 2011 ; Schön et al., 2018), de même que la pratique de la danse (Wu et al., 2021). Néanmoins, le niveau d'expertise nous semblant important à prendre en compte, nous avons choisi d'inclure des questions sur la pratique musicale et la pratique de danse dans le questionnaire socio-

démographique⁵⁸. Nous nous sommes également intéressée à l'expertise en jeux vidéo qui convoque elle aussi une maîtrise de la motricité fine par la manipulation de manettes de jeux (cf. Annexe 2, p. 239).

L'étude impliquant des tâches motrices, d'écoute et de lecture à haute voix, nous avons décidé de ne pas inclure les participants présentant des troubles de la motricité, de l'audition, de la vue (si non corrigés) ou du langage qui auraient pu rencontrer des difficultés liées aux tâches elles-mêmes. La précaution de ne pas inclure les participants présentant un trouble moteur permettait notamment de nous assurer que les tâches de l'Activité 1 (TMS et SMS) sont bien représentatives des capacités de traitement du rythme.

Les critères de sélection peuvent être synthétisés ainsi :

- Critères d'inclusion :
 - âge \geq 18 ans ;
 - langue maternelle : français ;
 - profil neurotypique ;
 - consentement à participer à l'étude.
- Critères de non inclusion :
 - présence de trouble visuel non corrigé ;
 - présence de trouble de l'audition, de la motricité ou du langage ;
 - musicien·ne ou danseur·se professionnel·le.

5.2 RECRUTEMENT

L'annonce de l'étude a été diffusée sur plusieurs sites toulousains *via* des tracts et les réseaux sociaux et universitaires. Un formulaire en ligne *via* l'outil LimeSurvey⁵⁹ (LimeSurvey Project Team, n.d.) permettait aux personnes intéressées de vérifier l'adéquation de leur profil selon les critères de sélection, avant de s'inscrire à un créneau de passation.

⁵⁸ Dans nos résultats, nous appellerons « musiciens » les participants ayant indiqué pratiquer ou avoir pratiqué une activité de musique et « danseurs » les participants ayant indiqué pratiquer ou avoir pratiqué une activité de danse.

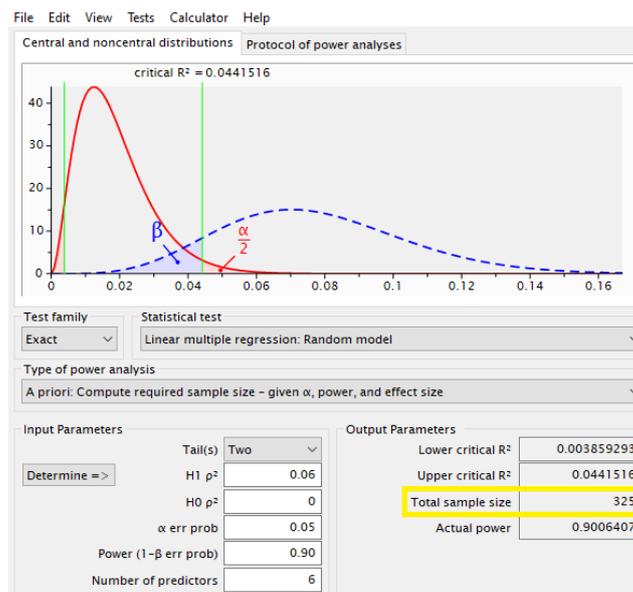
⁵⁹ Interface de création de questionnaire hébergée en ligne sur des serveurs de l'Université Toulouse 2.

5.3 ESTIMATION DU NOMBRE DE PARTICIPANTS

La taille de l'échantillon de participants a été estimée au préalable des passations expérimentales avec le logiciel G*Power (Faul et al., 2007) (voir **Figure 25**) à partir du plan expérimental de l'Activité 2, celui-ci étant le plan le plus complexe toutes tâches expérimentales confondues (pour un rappel des conditions expérimentales de l'Activité 2, se référer au **Tableau 2**, p. 94). Il en ressort que pour appliquer un modèle linéaire mixte à nos données avec une puissance de 0,90, en estimant un effet faible à modéré de l'hypothèse alternative ($H1 = 0,06^{60}$), soit un effet faible à modéré de nos six prédicteurs (effet de l'Engagement Moteur, du Type d'Amorce, de la Congruence Métrique et de leurs interactions doubles) sur notre variable dépendante (Temps d'Initiation de la parole), 325 observations sont nécessaires. Avec 8 observations par condition par participant, nous avons donc besoin d'une quarantaine de participants ($40 \times 8 = 320$ observations).

Figure 25 :

*Estimation du nombre de participants par le logiciel G*Power*



Note. Estimation à partir du plan expérimental de l'Activité 2.

5.4 DESCRIPTION DE L'ECHANTILLON

L'étude inclut 41 participants⁶¹ adultes neurotypiques francophones natifs (21 F ; âge moyen 31,3 ans ($\sigma = 9,44$)). Quant à leur profil d'expert, 27 participants ont indiqué pratiquer ou avoir pratiqué un instrument de musique à un niveau amateur ; 21 participants ont indiqué

⁶⁰ Effet faible : $\rho^2 = 0.02$; effet modéré : $\rho^2 = 0.13$; effet fort : $\rho^2 = 0.26$ (*G * Power 3.1 Manual*, 2021).

⁶¹ Les données d'un participant n'ont pas été retenues pour les analyses (cf. 6.3.2.1, p. 38).

pratiquer ou avoir pratiqué la danse à un niveau amateur et 28 ont indiqué jouer ou avoir joué aux jeux vidéo (cf. **Tableau 4**). La description de l'ensemble des caractéristiques socio-démographiques des participants est restituée en Annexe 3 (p. 241).

Tableau 4 :
Récapitulatif du profil socio-démographique des participants

| | Effectif | Âge | Pratique de la musique | Pratique de la danse | Pratique des jeux vidéo |
|--------|----------|---|--------------------------------|--------------------------------|--------------------------------|
| Femmes | 21 | âge moyen 28,05 ans ($\sigma = 6,64$) | Oui (n = 14) ; Non (n = 7) | Oui (n = 16) ; Non (n = 5) | Oui (n = 12) ; Non (n = 9) |
| Hommes | 20 | âge moyen 34,75 ans ($\sigma = 10,80$) | Oui (n = 13) ; Non (n = 7) | Oui (n = 5) ; Non (n = 15) | Oui (n = 16) ; Non (n = 4) |
| Total | 41 | âge moyen 31,3 ans (σ = 9,44) | Oui (n = 27) ; Non (n = 14) | Oui (n = 21) ; Non (n = 20) | Oui (n = 28) ; Non (n = 13) |

6 DONNEES RECUEILLIES

Les données recueillies concernent les données sociodémographiques du questionnaire (cf. 6.1), les données comportementales de l'Activité 1 (cf. 6.2, p. 114) et les données comportementales de l'Activité 2 (cf. 6.3, p. 119). Nous présenterons la nature et le pré-traitement des données opéré.

Concernant le traitement général des valeurs aberrantes, nous avons pris le parti d'enlever de l'analyse (1) des données au niveau individuel lorsque celles-ci s'écartent de plus ou moins deux écarts-types de la moyenne du participant (suppression des données sur une seule tâche expérimentale) et (2) des données au niveau du groupe de participants lorsque les performances du participant à une tâche semblent indiquer une non-conformité aux critères d'inclusion de l'étude (suppression des données de ce participant pour l'ensemble de l'étude). Nous avons ainsi exclu un participant (nous expliquons ce choix en partie 6.3.2.1, p. 120).

6.1 DONNEES DU QUESTIONNAIRE SOCIO-DEMOGRAPHIQUE

6.1.1 NATURE DES DONNEES

Les données personnelles recueillies relatives au profil des participants (informations générales) permettent de décrire la population de notre échantillon (cf. 5.4, p. 111) et éventuellement à éclairer les résultats des participants. Nous avons ainsi recueilli des données sur l'âge, le sexe, la latéralité, le plus haut diplôme obtenu, l'activité principale, le domaine d'activité, la ou les langues parlées, le lieu de résidence géographique, la pratique musicale, la pratique de la danse et la pratique des jeux vidéo (cf. Annexe 2, p. 239). Concernant la pratique musicale, de la danse et des jeux vidéo, les participants devaient renseigner par « Oui/Non » leur pratique et indiquer le nombre d'années de pratique et l'instrument de musique pratiqué.

6.1.2 PRE-TRAITEMENT DES DONNEES

Les questionnaires socio-démographiques étant au format papier, les réponses ont été ensuite rapportées manuellement dans l'outil LimeSurvey, afin de faciliter la création des tableaux de données nécessaires à l'analyse.

6.2 DONNEES DE L'ACTIVITE 1 : CARACTERISATION DES CAPACITES DE TRAITEMENT DU RYTHME

6.2.1 NATURE DES DONNEES ET VARIABLES DEPENDANTES

Les deux tâches de l'Activité 1 (Tempo Moteur Spontané et Synchronisation Sensori-Motrice) ont permis de recueillir des données comportementales relatives aux capacités de traitement du rythme en production, perception et synchronisation. Les données brutes recueillies correspondent aux chronométrages associés aux frappes du doigt du participant sur une touche du boîtier de réponse, et sont fournies par le logiciel PsychoPy dans un classeur de sortie.

Plus précisément, lors de la tâche de TMS, nous avons recueilli deux séquences de chronométrages, puisque la collecte se fait en deux blocs expérimentaux. Pour la tâche de SMS, nous avons recueilli six séquences de chronométrages, relatives aux trois tempi administrés aux participants en deux blocs expérimentaux.

A l'issue de la passation expérimentale de l'Activité 1, le logiciel PsychoPy émettait un classeur Excel pour chaque participant contenant l'identifiant du participant ; les chronométrages associés à la tâche de TMS ; les chronométrages associés à la tâche de SMS ainsi que l'ordre d'enchaînement des tempi (ceux-ci étant aléatorisés entre les participants).

Nous avons ensuite transformé ces données brutes afin d'obtenir nos différentes variables dépendantes (nous détaillons les étapes de transformation des données dans la partie suivante 6.2.2, p. 115). Pour rappel, pour la tâche de TMS, les variables dépendantes sont le TMS moyen et la stabilité du TMS ; pour la tâche de SMS, les variables dépendantes sont la stabilité de la SMS et la précision de la SMS. Pour la mesure du TMS moyen, nous avons calculé la moyenne des intervalles temporels inter-battements (IBI) de chaque participant, un IBI correspondant à la différence temporelle entre les chronométrages associés à deux frappes du doigt. Pour les mesures de stabilité du TMS, de stabilité de la SMS et de précision de la SMS, nous avons calculé la longueur (stabilité) et la direction (précision) du vecteur représentatif de la performance moyenne du participant dans chacune de ces deux tâches, à l'aide des statistiques circulaires que nous développons dans la partie suivante.

6.2.2 PRE-TRAITEMENT DES DONNEES : STATISTIQUES CIRCULAIRES

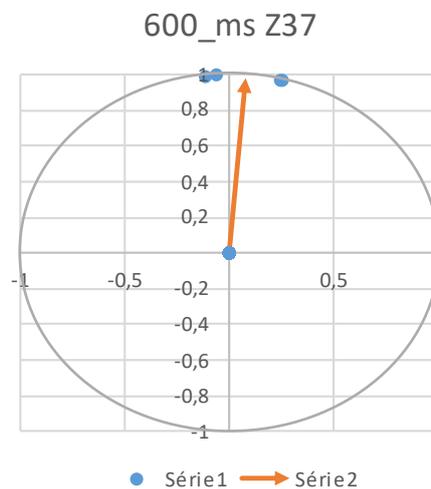
Afin de mesurer la stabilité du TMS ainsi que la stabilité et la précision de la SMS des participants, les chronométrages (en ms) recueillis pour les tâches de TMS et SMS ont été transformés en intervalles temporels inter-battements (IBI) (en ms) puis en mesures d'angles (en degrés) et enfin en vecteurs circulaires (nous détaillons ces étapes en suivant).

Dans la tâche de TMS, les valeurs d'IBI (en ms) s'écartant de ± 2 écarts-types de la moyenne (pour chaque participant) ont été supprimées, soit une suppression de 106 valeurs aberrantes sur 2.560⁶² observations initialement (4,14 %). Dans la tâche de SMS, les valeurs d'IBI (en degrés) s'écartant de ± 2 écarts-types de la moyenne des asynchronies par rapport au tempo du métronome (pour chaque participant) ont été supprimées, soit une suppression de 711 valeurs aberrantes sur 22.893⁶³ observations initialement (3,11 %).

Pour les deux tâches de l'Activité 1 (TMS et SMS), nous avons choisi de passer par la méthode des statistiques circulaires (Kirschner & Tomasello, 2009), plus adaptées que les statistiques linéaires face à la nature continue et cyclique des données temporelles des battements de métronome et des frappes du doigt du participant (pour une représentation graphique, cf. **Figure 26** et **Tableau 5**).

Figure 26 :

Représentation circulaire de 4 frappes pour un extrait de la tâche de SMS



Note. Représentation circulaire de 4 frappes pour un extrait de la tâche de SMS du participant Z37 (ISI métronome = 600 ms). En bleu sur le cercle : la dispersion des frappes du participant ; flèche en orange : le vecteur représentatif de la performance moyenne du participant, créé à partir du **Tableau 5**. La longueur du vecteur (\bar{R}) reflète la stabilité moyenne du participant ; l'angle du vecteur ($\bar{\theta}$) par rapport « nord » du cercle représente l'asynchronie moyenne du participant.

⁶² 32 IBI * 2 blocs * 40 participants (un participant ayant été exclu de l'analyse, cf. 6.3.2.1, p. 120).

⁶³ IBI recueillis sur 40,5 secondes par tempo * 2 blocs * 40 participants (un participant ayant été exclu de l'analyse, cf. 6.3.2.1, p. 120).

Tableau 5 :
Exemple de 4 frappes extraites de la SMS au tempo 600 ms

| ID | ISI métronome (ms) | Chrono frappes (ms) | IBI (ms) | Modulo | Angle (°) | Sinus (absc.) | Cosinus (ordon.) | $\bar{\theta}$ (°) | \bar{R} | p -value (test de Rayleigh) |
|-----|--------------------------|---------------------------|----------|--------|--------------|------------------|---------------------|--------------------|-----------|-------------------------------------|
| Z37 | 600 | 1189,89 | 1189,89 | 588,89 | -6,672 | -0,116 | 0,993 | 4,861 | 0,985 | 0,009 |
| Z37 | 600 | 1824,55 | 635,67 | 24,55 | 14,728 | 0,254 | 0,967 | | | |
| Z37 | 600 | 2424,3 | 599,75 | 24,3 | 14,58 | 0,252 | 0,968 | | | |
| Z37 | 600 | 2994,67 | 570,37 | 594,67 | -3,2 | -0,056 | 0,998 | | | |

Note. Exemple de 4 frappes extraites du participant Z37 pour la tâche de SMS (ISI métronome = 600 ms), placées sur la **Figure 26**. Dans cet exemple, la valeur du test de Raleigh est inférieure à 0,05, indiquant une distribution des frappes du participant non aléatoire (réussite de la synchronisation au tempo de métronome 600 ms).

Concernant la SMS, les statistiques linéaires se basent sur l’alignement temporel entre les battements du métronome et les frappes du participant, supposant une correspondance biunivoque (« un à un ») entre un battement et une frappe. Cette correspondance peut s’avérer problématique lorsque le participant se décale, qu’il effectue une frappe de trop ou de moins. L’approche circulaire permet alors de contourner ce genre de situations : on ne cherche plus à aligner chaque frappe à chaque battement de manière linéaire. La séquence de battements du stimulus (métronome) étant isochrone, cette séquence elle-même devient la nouvelle échelle, non plus linéaire et infinie, mais circulaire et fermée, comme une boucle.

Ce procédé se retrouve notamment dans la manière traditionnelle de compter les heures de la journée, représentée graphiquement par un cadran horaire circulaire, dont l’échelle est l’heure. Le principe que nous appliquons pour les statistiques circulaires est le même : notre échelle circulaire n’est plus l’horloge (une heure = 360°) mais l’intervalle inter-stimuli (ISI⁶⁴) du métronome (par ex. 600 ms = 360°) dans la tâche de SMS ou la moyenne des intervalles inter-battements (IBI) dans la tâche de TMS. Chaque frappe du participant est alors placée sur le cercle trigonométrique en fonction de son rapport à cette échelle. Les statistiques circulaires nous permettent ainsi de rendre compte de la dispersion (renvoyant à l’asynchronie) des frappes ainsi que de la performance moyenne (la variance et la moyenne de ces asynchronies).

Pour passer des IBI en ms à un vecteur circulaire, plusieurs opérations doivent être conduites (cf. **Tableau 5** et **Figure 26**) :

1. établissement de l’échelle circulaire : l’intervalle inter-stimuli (ISI) du métronome ou le TMS moyen du participant est converti en degrés (par exemple, 600 ms = 360°), qui vont nous permettre de travailler sur le cercle trigonométrique (échelle circulaire = 360° = 1 tour de cercle) ;

⁶⁴ De l’anglais *Inter-Stimulus Interval*.

2. placement des frappes du participant sur le cercle :
 - a. calcul du modulo de chaque chronométrage (en ms) du participant en fonction de l'ISI du métronome ou du TMS moyen. Le modulo nous donne le reste de la division entière du chronométrage par l'ISI, ce reste illustrant l'asynchronie (positive) de chaque frappe par rapport à l'échelle temporelle établie ;
 - b. conversion des modulo (asynchronies en ms) en degrés afin de pouvoir les placer sur le cercle ayant pour échelle l'ISI. Chaque frappe du participant est donc associée à un angle, représentatif du décalage entre le battement entendu (placé à 0°) et cette frappe, la valeur de l'angle pouvant aller de 0° à 360° . Cependant afin de rendre compte de la direction de l'asynchronie (battement de doigt du participant en avance ou en retard par rapport au métronome), les battements associés à un angle entre 180° et 360° sont convertis de sorte à ce que leur valeur se situe entre -180° et 0° . Si le battement de doigt du participant arrive entre 0° et 180° , cela signifie qu'il est en retard par rapport au battement du métronome. Si le battement de doigt du participant arrive entre -180° et 0° , cela signifie qu'il est en avance par rapport au battement du métronome ;
 - c. calcul du cosinus et du sinus de chaque angle afin de les placer sur un cercle de rayon 1, intégré dans un repère ayant pour coordonnées x (sinus) et y (cosinus) ;
 - d. les frappes peuvent ensuite être placées sur un cercle (échelle = ISI), faisant état de la dispersion des frappes du participant.
3. A partir de la dispersion (angles représentatifs de l'asynchronie de chaque frappe par rapport à l'ISI du métronome ou par rapport au TMS moyen), il est possible de calculer la performance moyenne du participant en construisant un vecteur moyen \bar{R} . Ce vecteur comporte 2 composantes :
 - a. une direction moyenne $\bar{\theta}$ (correspondant à la valeur et au signe de l'angle moyen, soit l'asynchronie moyenne) ;
 - b. une longueur résultante moyenne \bar{R} (correspondant à l'inverse de la variance circulaire des asynchronies, comprise entre 0 et 1). Plus un participant est stable, c'est-à-dire qu'il a la même fréquence que l'ISI du stimulus, plus la longueur du vecteur se rapproche de 1. La significativité de la longueur résultante moyenne \bar{R} est mesurée avec le test de Rayleigh, dont le seuil de significativité p est fixé à 0,05. Un résultat $\leq 0,05$ illustre une différence significative de la répartition des frappes du participant par rapport à une distribution aléatoire (uniforme) sur le pourtour du cercle, et donc la réussite de la synchronisation.

Le vecteur nous donne ainsi deux informations sur la performance moyenne du participant, que nous utilisons en variables dépendantes pour les tâches de TMS et SMS : la précision (*via* l'asynchronie moyenne, en degrés) du participant par sa direction moyenne $\bar{\theta}$ et la stabilité du participant par sa longueur résultante moyenne \bar{R} (comprise entre 0 et 1).

Le logiciel Excel (Microsoft Corporation, 2021) a été utilisé pour calculer les composantes des vecteurs (direction/angle et longueur résultante) et le package « circular » de RStudio (R Core Team, 2021) a été utilisé pour calculer la significativité du test de Rayleigh.

6.3 DONNEES DE L'ACTIVITE 2 : AMORÇAGE RYTHMIQUE DE LA PAROLE

6.3.1 NATURE DES DONNEES

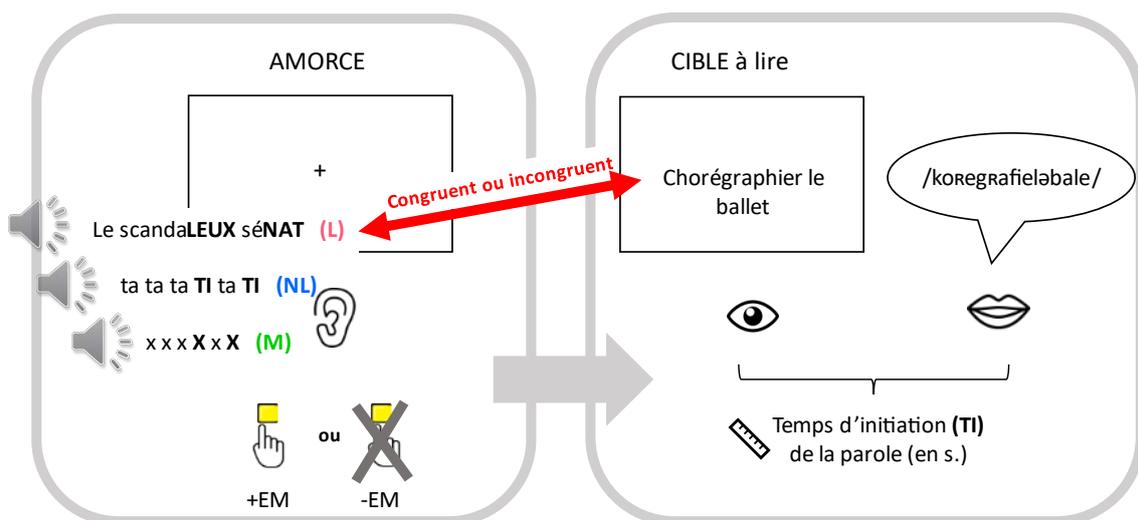
Lors de l'Activité 2, nous avons recueilli un enregistrement de parole pour chaque syntagme cible lu lors de la phase expérimentale proprement dite, soit un total de 96 syntagmes enregistrés par participant. Nous avons ainsi recueilli 3936 enregistrements de syntagmes (96*41).

A l'issue de chaque passation expérimentale (donc pour chaque participant) pour l'Activité 2, le logiciel PsychoPy émettait 96 fichiers audio correspondant à chacun des syntagmes lus ainsi qu'un classeur Excel contenant l'identifiant du participant, l'ordre d'enchaînement des amorces et des cibles et l'ordre d'enchaînement des blocs expérimentaux (pour un rappel des possibilités d'enchaînement des blocs expérimentaux, se rapporter à la **Figure 23** p. 101 et à la partie 3.2.4 p. 102).

A partir des enregistrements de parole, nous avons relevé deux variables dépendantes : le nombre d'erreurs de lecture et le temps d'initiation (TI) de la parole, ce dernier correspondant à la durée du silence entre l'apparition du syntagme cible à lire et le début du syntagme lu prononcé par le participant (cf. **Figure 27**).

Figure 27 :

Illustration de la mesure du temps d'initiation (TI) de la parole



Note. Le Temps d'Initiation (TI) de la parole correspond à la durée entre le moment d'apparition à l'écran du syntagme cible à lire et le début de l'élocution par le participant.

6.3.2 PRE-TRAITEMENT DES DONNEES

6.3.2.1 NOMBRE D'ERREURS DE LECTURE

Pour comptabiliser le nombre d'erreurs de lecture, nous avons relevé les syntagmes contenant des répétitions ou reformulations (par exemple, le syntagme « Ce sont des vieux volets » lu « Ce sont de vieux volets ce sont des vieux volets »); des substitutions grammaticales (par exemple, le syntagme « il ne met plus le gilet » lu « il ne met plus de gilet »), lexicales (par exemple, le syntagme « Elle en a mis une pincée » lu « Elle en a pris une pincée »), phonétiques (par exemple, le syntagme « Je prendrai le pari » lu « Je brendrai le pari »), phonologiques (par exemple, le syntagme « La cuisson du civet » lu « La fuisson du civet »), de genre (par exemple, le syntagme « Elle a commis un délit » lu « Il a commis un délit ») et dérivationnelles (par exemple, le syntagme « Chorégrapheur le ballet » lu « Chorégraphie le ballet »); les syntagmes avec omissions phonologiques (par exemple, le syntagme « Le spectacle est le jeudi » lu « e spectacle est le jeudi », morphologiques (par exemple, le syntagme « Il a pris le maquis » lu « Il pris le maquis ») ou lus partiellement (par exemple, le syntagme « Mon cardigan en pilou » lu « Mon cardigan en »); les syntagmes avec déplacement phonologique (par exemple, le syntagme « L'éblouissante chaussée » lu « L'éboulissante chaussée »); les syntagmes avec ajouts de phones (par exemple, le syntagme « Tous les glaçons ont fondu » lu « Tous les glaçons z'ont fondu ») et lexicaux (par exemple, le syntagme « Elle a un grand chalet » lu « Elle a fait un grand chalet »). Un syntagme comportant au moins une erreur était coté « 1 » et un syntagme sans erreur « 0 ». Toutes les productions cibles des participants ont été transcrites dans le logiciel Praat (version 6.3.01) (Boersma & Weenink, 2022).

Nous avons fait le choix d'exclure un participant présentant un nombre d'erreurs plus élevé que le reste du groupe (21 erreurs, par rapport à une médiane du groupe à 3 erreurs, cf. **Tableau 6** et **Figure 28**), ce qui nous paraissait refléter une difficulté langagière et donc contrevenant aux critères d'inclusion de l'étude (cf. 5.1, p. 109). Nous avons donc également enlevé de l'analyse les données de ce participant concernant les mesures du TI de la parole (Activité 2) et de TMS et SMS (Activité 1).

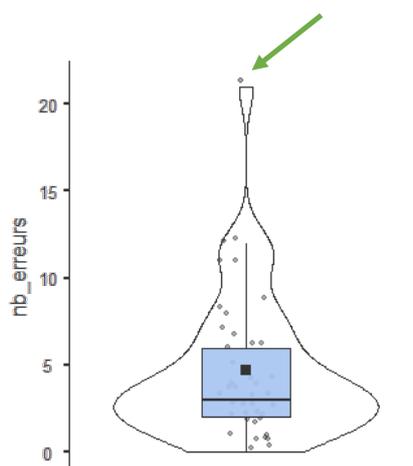
Tableau 6 :
Valeurs descriptives du nombre d'erreurs de lecture des 41 participants

Statistiques descriptives

| | N | Manquants | Moyenne | Médiane | Ecart-type | Minimum | Maximum | Shapiro-Wilk | |
|------------|----|-----------|---------|---------|------------|---------|---------|--------------|--------|
| | | | | | | | | W | p |
| nb_erreurs | 41 | 0 | 4.71 | 3 | 4.17 | 0 | 21 | 0.825 | < .001 |

Note. Un $p < 0,05$ au test de Shapiro-Wilk indique une distribution non paramétrique des données. Nous pouvons remarquer que le plus grand nombre d'erreurs (colonne Maximum) est de 21, qui correspond au participant dont nous avons ensuite écarté les données de l'analyse.

Figure 28 :
Représentation graphique du nombre d'erreurs de lecture des participants



Note. $N = 41$. Nous pouvons remarquer qu'un participant se distingue du reste du groupe en présentant un nombre d'erreurs plus élevé (sous la flèche verte), qui correspond au participant dont nous avons ensuite écarté les données de l'analyse.

6.3.2.2 TEMPS D'INITIATION DE LA PAROLE

Pour recueillir le Temps d'Initiation (TI) de la parole, nous avons opté pour une segmentation manuelle avec le logiciel Praat assistée par le plug-in EasyAlign (Goldman, 2011) (un exemple de segmentation est présenté **Figure 29**). La segmentation s'est faite principalement à l'aide d'indices visuels du spectrogramme (cf. **Tableau 7**).

Les valeurs de Temps d'Initiation s'écartant de ± 2 écarts-types de la moyenne (pour chaque participant) ont été enlevées de l'analyse⁶⁵. Sur 40 participants, cela correspond à la

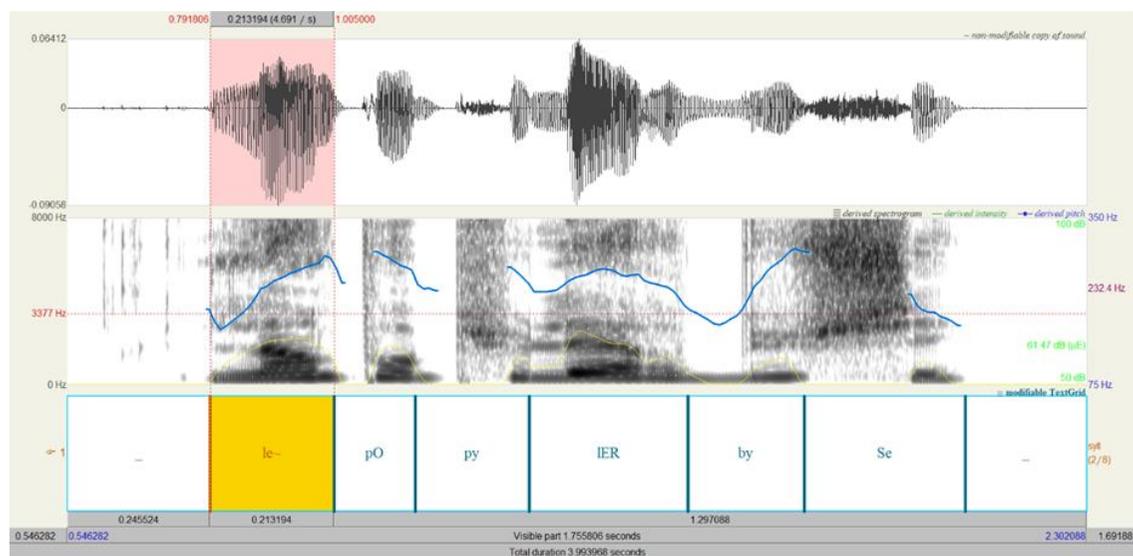
⁶⁵ D'après les recommandations de Pitt et Samuel (1990) concernant les valeurs aberrantes.

suppression de 186 valeurs aberrantes sur 3840 observations initialement, soit une suppression de 4,84 % des observations.

Tableau 7 :
Indices de segmentation pour recueillir le Temps d'Initiation de la parole (TI).

| | Voyelles | Consonnes sonantes (nasales, liquides, glides) | Consonnes obstruantes | | | |
|-------------------------|--|--|-----------------------|--------------------|--|---|
| | | | Plosives sourdes | Plosives sonores | Fricatives sourdes | Fricatives sonores |
| Indices de segmentation | concomittance de la courbe de F0 et d'intensité ; présence richesse spectrale importante | concomittance de la courbe de F0 et d'intensité ; présence richesse spectrale importante | barre d'explosion | barre de voisement | présence de bruit (large bande fréquentielle d'énergie) dans les hautes fréquences | concomittance bruit dans les hautes fréquences + barre de voisement |
| + indices acoustiques | | | | | | |

Figure 29 :
Exemple d'une segmentation en syllabes du syntagme « L'impopulaire bûcher »



Note. L'emplacement de frontière gauche de la syllabe initiale (colorée) correspond au TI de la parole. Ligne bleue : courbe de f0 ; ligne jaune : courbe d'intensité

7 DEMARCHE STATISTIQUE

Dans nos analyses, nous démarrons à chaque fois par une exploration descriptive des données recueillies. Dans cette optique, les statistiques descriptives sont effectuées sur données agrégées (moyennées par participant et par condition expérimentale) afin d'observer les premières tendances. Ensuite, nous avons modélisé la relation entre les variables manipulées dans l'expérience (variables indépendantes – VI) et les variables dépendantes (VD) observées, à l'aide de modèles de régression (cf. 7.1). Lorsque la VD est quantitative, nous avons appliqué un modèle linéaire mixte tandis que lorsque la VD est binomiale⁶⁶, nous avons appliqué un modèle linéaire généralisé à effets mixtes binomial.

Les analyses descriptives et inférentielles sont effectuées à l'aide du plug-in Gamlj3 (Gallucci, 2019) du logiciel Jamovi (The jamovi Project, 2023).

7.1 MODELE MIXTE

Le principe du modèle de régression **linéaire** est de formuler mathématiquement la relation existant entre deux variables, une VI (qualitative ou continue) et une VD (continue), et ainsi de prédire les valeurs de la variable dépendante en fonction des niveaux de la variable indépendante⁶⁷. La relation entre une ou plusieurs variables indépendantes (par exemple, l'engagement moteur du participant) et une variable dépendante (par exemple, le temps d'initiation de la parole) peut être modélisée par une équation mathématique de droite de type $y = ax + b$. Dans cette équation, y est la valeur prédite (VD) et x la variable prédictrice (VI manipulée dans l'expérience). La variable prédictrice (ou 'prédicteur') peut être un facteur (variable catégorielle) ou une covariable continue (variable numérique). Dans l'équation, la valeur de y varie en fonction de la valeur de l'ordonnée à l'origine b et de la force (valeur de a) et de la direction (signe de a) de la relation entre y et x . La valeur et le signe de a correspondent au coefficient de régression de la variable prédictrice x . La régression peut ensuite être qualifiée de simple ou multiple, selon que le modèle intègre un seul (simple) ou plusieurs (multiple) prédicteurs.

Le modèle de régression **binomiale** s'appuie sur le principe de la modélisation linéaire, mais modélise une transformation de la VD, celle-ci ne pouvant prendre que deux valeurs (0 ou

⁶⁶ Une variable binomiale ne comporte que deux valeurs (par exemple, « 0 » pour « pas d'erreur de lecture » et « 1 » pour « présence d'au moins une erreur de lecture »).

⁶⁷ Cf. le manuel de statistiques en sciences du langage de Solier et al. (2023).

⁶⁸). L'équation de la droite de régression devient donc $f(y) = ax + b$, où $f(y)$ représente une fonction de lien *logit*. Le modèle de régression binomial établit une échelle de probabilité allant de 0 à 1, et prédit la probabilité pour y de prendre la valeur 1⁶⁹. Par exemple, si « 0 » correspond à « aucune erreur de lecture produite dans le syntagme » et « 1 » à « au moins une erreur de lecture produite dans le syntagme », le modèle prédit la probabilité qu'un syntagme produit présente au moins une erreur de lecture (« 1 ») en fonction des paramètres du modèle (les VI).

Le modèle **mixte**, qu'il soit linéaire ou binomial, permet d'intégrer au modèle des effets aléatoires en plus des effets fixes (effets des VI), afin de prendre en compte la variabilité due à des facteurs non mesurés (les effets aléatoires). Les effets aléatoires renvoient à des variables qualitatives échantillonnées de manière aléatoire dans leur groupe respectif lors de l'expérience, et se manifestent par des observations dépendantes entre elles (par exemple : les observations associées à un même participant ou à un même item cible). Les effets fixes sont donc les effets que l'on cherche à généraliser à une population plus importante tandis que les effets aléatoires représentent la variabilité inhérente à nos données, la déviation par rapport à l'effet des variables prédictives.

Dans la réalité des faits, le modèle ne prédit pas parfaitement chacune des valeurs observées de la variable dépendante, en raison du caractère naturel des données comportementales recueillies. Afin de sélectionner le modèle le plus ajusté à nos données, nous avons suivi la procédure « *top-down* » recommandée par Zuur et al. (2009, p. 121-122 et p. 540-541) concernant la comparaison de modèles mixtes : (1) construire un modèle saturé comportant un maximum de paramètres (effets fixes et interactions des effets fixes) ; (2) à partir du modèle saturé et en utilisant l'estimation restreinte du maximum de vraisemblance (REML⁷⁰), comparer la pertinence des différents effets aléatoires possibles pour trouver la structure aléatoire optimale ; (3) à partir du modèle saturé avec une structure aléatoire optimale et en utilisant cette fois l'estimation (non restreinte) du maximum de vraisemblance (ML⁷¹), comparer la pertinence des différents effets fixes pour parvenir à la structure fixe optimale ; (4) présenter le modèle final avec l'estimation REML⁷².

⁶⁸ Par exemple, dans l'Activité 2 d'amorçage rythmique de la parole, les syntagmes comportant au moins une erreur sont cotés « 1 » et ceux sans erreur sont cotés « 0 ».

⁶⁹ Cf. le manuel de statistiques de Zuur et al. (2009).

⁷⁰ *REstricted Maximum Likelihood*.

⁷¹ *Maximum Likelihood*.

⁷² Le module GAMLj de Jamovi permet de faire la distinction entre les estimations ML et REML pour le modèle linéaire mixte mais pas pour le modèle linéaire généralisé à effets mixtes binomial. Nous n'avons donc pas fait de comparaison des effets fixes pour les modèles binomiaux.

Une fois que le modèle (linéaire ou binomial) (notre équation modélisant la relation entre y et x) est considéré comme adéquat, on peut alors observer l'effet de la ou des variables indépendantes sur la variable dépendante avec un test omnibus (test F ou chi-2) de comparaison des variances ou probabilités estimées.

7.1.1 APPLICATION DES MODELES MIXTES AUX DONNEES RECUEILLIES

Dans les analyses inférentielles effectuées, nous avons cherché d'une part à estimer la part des variables indépendantes manipulées (VI « d'intérêt ») afin de discuter nos hypothèses (cf. *Hypothèse 1* à *Hypothèse 6* p. 78) et d'autre part, nous avons voulu contrôler cet effet à l'aide de variables additionnelles (VI « contrôles »). L'ajout de variables contrôles dans le modèle (par exemple, les caractéristiques socio-démographiques) permet d'observer si la variance observée dans la VD est bien liée aux VI d'intérêt ou bien si l'effet de ces dernières est en fait modulé par d'autres paramètres (les VI contrôles) (cf. Cohen et al., 2003, p. 608). L'interprétation des modèles comprenant des variables contrôles est néanmoins à faire avec précaution, en raison du caractère non manipulé de ces variables et du risque que le modèle soit très bien ajusté à notre échantillon mais de fait moins généralisable et donc moins bien adapté pour une population différente par exemple du point de vue socio-démographique. Concernant le codage des variables, nous avons utilisé le codage de base de Jamovi, soit un codage « simple » pour les VI catégorielles et un codage « centered » pour les VI numériques.

Nous avons considéré comme variables contrôles l'âge, le nombre d'années de pratique de la musique et de la danse, le sexe (H/F), la pratique de la musique, de la danse et des jeux vidéo (Oui/Non). En effet, par exemple, selon Drake, Jones, et al. (2000), l'avancée de l'âge et la pratique musicale sont associées à un ralentissement de la période de référence et à une augmentation de la précision de la SMS. Cela se traduirait par un ralentissement du TMS, du TI de la parole et du tempo auquel la SMS est la plus stable et la plus précise ainsi que par une stabilité et une précision plus importante dans l'Activité 1, avec l'avancée de l'âge et la pratique musicale. De plus, l'expertise rythmique amenée par la pratique de la musique et de la danse pourrait se manifester par une sensibilité accrue à la congruence métrique dans la tâche de lecture après amorçage rythmique (Activité 2). La pratique des jeux vidéo pourrait également impacter la précision rythmique motrice dans les tâches de TMS et SMS.

Dans l'Activité 1, pour investiguer l'*Hypothèse 1* et l'*Hypothèse 6*, nous avons construit et appliqué deux modèles linéaires mixtes pour la tâche de caractérisation du Tempo Moteur

Spontané et quatre modèles linéaires mixtes pour la tâche de caractérisation de la Synchronisation Sensori-Motrice. Dans l'Activité 2 (tâche d'amorçage rythmique de la parole), pour investiguer les Hypothèse 2 à 5 nous avons construit et appliqué un modèle linéaire généralisé à effets mixtes binomiaux et trois modèles linéaires mixtes.

Nous n'avons pas manipulé de VI d'intérêt dans la tâche de TMS, dont le but était de recueillir deux VD : le TMS moyen et la stabilité du TMS. Les deux modèles linéaires mixtes construits et appliqués aux données ont permis d'observer la relation entre d'une part les caractéristiques socio-démographiques des participants et d'autre part le TMS moyen ou la stabilité du TMS.

Dans la tâche de SMS, nous avons manipulé le tempo du métronome (trois tempi) et nous avons recueilli deux VD : la stabilité et la précision de la SMS. Deux modèles linéaires mixtes ont été construits pour chacune des deux VD : un modèle estimant uniquement l'effet de la VI d'intérêt (le tempo) sur la VD (stabilité / précision de la SMS) et un second modèle estimant l'effet de la VI d'intérêt ainsi que de VI contrôles socio-démographiques.

Dans l'Activité 2, pour investiguer nos Hypothèse 2 à 5, nous avons manipulé trois VI (l'Engagement Moteur du participant – EM, le Type d'Amorce – TA et la Congruence Métrique – CM) et avons recueilli deux VD (le nombre d'erreurs de lecture et le Temps d'Initiation de la parole – TI). Nous avons construit et appliqué un modèle linéaire généralisé à effets mixtes binomial pour estimer l'effet des trois VI (EM, TA et CM) sur les erreurs de lecture. Nous avons également construit et appliqué trois modèles linéaires mixtes pour estimer l'effet des trois VI sur le TI de la parole. Le premier modèle estime uniquement l'effet des VI d'intérêt (EM, TA et CM) sur la VD (TI de la parole), le deuxième modèle estime l'effet des VI d'intérêt ainsi que de VI contrôles relatives au profil rythmique et socio-démographique et le troisième modèle estime l'effet des VI d'intérêt ainsi que l'effet du groupe métrique de l'amorce.

7.1.2 PRESENTATION DES RESULTATS

Lorsque nous présentons les résultats des modèles mixtes, nous annonçons la contribution individuelle des effets fixes par la valeur de la *p-value* résultant du test omnibus (test F pour le modèle linéaire et test du Chi-2 pour le modèle logistique) ainsi que par la valeur

du coefficient de la régression (*Estimate* pour les covariables ou *Difference* pour les facteurs) (pour le modèle linéaire) ou la valeur de l'*odd-ratio* (*OR* ou $Exp(B)$) (pour le modèle logistique).

IV. RESULTATS

Dans cette section, nous présentons les résultats obtenus dans les différentes tâches expérimentales (cf. 8, p. 133 et 9, p. 169) ainsi que les comparaisons entre tâches (cf. 10, p. 187). Pour rappel, afin d'observer **l'existence d'une période de référence supramodale, l'effet de l'amorçage rythmique sur la parole et les liens entre traitement du rythme et du langage**, nous avons posé six hypothèses (cf. *Hypothèse 1 à 6*, p. 78), investiguées dans **deux activités expérimentales**.

Pour commencer, nous présentons les résultats de l'Activité 1 (cf. 8, p. 133), soit les résultats des tâches de Tempo Moteur Spontané (TMS) et de Synchronisation Sensori-Motrice (SMS).

Le tempo le plus stable et le plus précis des participants dans la tâche de SMS sera mis en regard de leur TMS moyen afin d'investiguer l'*Hypothèse 1.1*. En effet, nous supposons que le tempo de la SMS le plus stable et/ou le plus précis des participants correspond au tempo (parmi les trois proposés : 225, 600 et 900 ms) se rapprochant le plus de leur TMS moyen. Dans la troisième partie de cette section des résultats (partie 10, p. 187), les mesures de TMS moyen et de tempo de la SMS le plus stable et/ou le plus précis (Activité 1) sont mises en regard entre elles (cf. *Hypothèse 1.1*) et en regard du Temps d'Initiation (TI) de la parole recueilli dans l'Activité 2 afin d'investiguer l'*Hypothèse 1.2*. En effet, selon l'*Hypothèse 1.2*, nous supposons que le TI de la parole moyen (Activité 2) est corrélé au TMS moyen (Activité 1).

Les tâches de TMS et SMS nous permettent également de recueillir des données à propos des capacités de traitement du rythme, à savoir la stabilité du TMS, et la stabilité et la précision de la SMS. Ces mesures sont mises en regard des capacités de traitement du langage (nombre d'erreurs de lecture et TI de la parole) recueillies dans l'Activité 2 afin d'investiguer l'*Hypothèse 6* dans la partie 10 (p. 187). En effet, nous supposons que les participants les plus stables et les plus précis lors de l'Activité 1 (TMS et SMS) présentent moins d'erreurs de lecture et un TI de la parole plus court lors de l'Activité 2.

La deuxième partie est consacrée à la présentation des résultats de l'Activité 2 (cf. 9, p. 169), au cœur de notre étude, soit les résultats de la tâche de lecture à haute voix après amorçage rythmique, dans laquelle nous avons recueilli deux mesures : le nombre d'erreurs de lecture et le Temps d'Initiation (TI) de la parole.

Ces deux mesures nous ont permis d'observer l'effet de différences d'amorçage rythmique sur la facilitation de la production de parole (cf. *Hypothèse 2 à 5*, p. 79). En effet, nous supposons que :

- l'amorçage rythmique **avec engagement moteur** des participants a un effet **positif** sur la production de parole par rapport à l'amorçage rythmique auditif sans engagement moteur. Cela se manifeste par un nombre d'erreurs de lecture plus faible et un temps d'initiation de la parole plus court en condition d'amorçage rythmique avec engagement moteur (*Hypothèse 2*) ;
- l'amorçage **vocal langagier** a un effet **positif** pour la production de la parole par rapport à l'amorçage vocal non langagier, ce dernier lui-même plus facilitant que l'amorçage musical. Cela se manifeste par un nombre d'erreurs de lecture plus faible et un temps d'initiation de la parole plus court en condition d'amorçage vocal langagier par rapport à l'amorçage rythmique vocal non langagier et l'amorçage rythmique musical, ce dernier étant associé à un nombre d'erreurs de lecture le plus élevé et un temps d'initiation de la parole le plus long (*Hypothèse 3*) ;
- l'amorçage **métriquement congruent** (structure métrique sous-jacente similaire entre le stimulus d'amorce rythmique et le syntagme cible à lire) a un effet **positif** pour la production de parole par rapport à l'amorçage incongruent. Cela se manifeste par un nombre d'erreurs de lecture plus faible et un temps d'initiation de la parole plus court en condition d'amorçage congruent par rapport à l'amorçage incongruent (*Hypothèse 4*) ;
- l'effet de la **congruence métrique** de l'amorçage est **renforcé par l'engagement moteur** des participants. Cela se manifeste par un nombre d'erreurs de lecture le plus faible et un temps d'initiation de la parole le plus court en condition d'amorçage congruent avec engagement moteur, et par un nombre d'erreurs de lecture le plus important et par un temps d'initiation de la parole le plus long en condition d'amorçage incongruent avec engagement moteur (*Hypothèse 5*).

Dans une dernière partie enfin (cf. 10, p. 187), nous mettons en regard les résultats de l'Activité 1 et de l'Activité 2, afin de tester l'*Hypothèse 1* et l'*Hypothèse 6*.

8 ACTIVITE 1 : CARACTERISATION DES CAPACITES DE TRAITEMENT DU RYTHME

Dans une première activité, nous avons recueilli des mesures représentatives des capacités de traitement du rythme, dans le but de les mettre en regard des capacités de traitement du langage observées dans l'Activité 2 dans un second temps (partie 10, p. 187).

Les capacités de traitement du rythme ont été décrites dans une tâche de production du Tempo Moteur Spontané (TMS) et une tâche de Synchronisation Sensori-Motrice (SMS⁷³). La tâche de TMS nous a permis de décrire les capacités de traitement du rythme en production et renseigne sur le tempo de préférence moteur des participants ; la tâche de SMS nous a permis de décrire les capacités de traitement du rythme en perception et synchronisation à différents tempi de métronome et renseigne sur le tempo auquel la synchronisation est la plus stable et la plus précise.

L'Activité 1 nous permet de tester l'*Hypothèse 1*, concernant l'existence d'une période de référence supramodale. Le Tempo Moteur Spontané (TMS) moyen reflète la période de référence. Nous supposons que le tempo auquel la Synchronisation Sensori-Motrice (SMS) est la plus stable et/ou précise d'un participant sera le tempo le plus proche (parmi les trois proposés : 225, 600 et 900 ms) de son TMS moyen. Le TMS moyen et le tempo de SMS le plus stable et/ou précis seront ensuite mis en regard du Temps d'Initiation (TI) de la parole mesuré dans l'Activité 2 dans la partie 10 (p. 187).

L'activité 1 nous permet également de tester l'*Hypothèse 6*, concernant l'existence de traitements partagés pour le rythme et le langage : la mesure de stabilité du TMS et les mesures de stabilité et précision de la SMS, qui reflètent les capacités de traitement du rythme, seront mises en regard du nombre d'erreurs de lecture et du TI de la parole, qui reflètent les capacités de traitement du langage, dans la partie 10 (p. 187).

Nous commencerons cette partie-ci avec la présentation des résultats de la tâche de TMS (cf. 8.1), puis celle de SMS (cf. 8.2, p. 145). Ces deux tâches seront ensuite mises en relation entre elles et avec l'Activité 2 afin d'investiguer l'*Hypothèse 1* et l'*Hypothèse 6* (cf. partie 10, p. 187).

⁷³ De l'anglais *Sensori-Motor Synchronization*.

8.1 TACHE DE TEMPO MOTEUR SPONTANE

Dans la tâche de Tempo Moteur Spontané (TMS), les participants avaient pour consigne de battre la pulsation en appuyant avec leur doigt sur une touche à une vitesse qui leur semblait agréable, naturelle. La mesure du TMS s'effectuait sur deux blocs expérimentaux de 33 battements chacun, nous avons ainsi recueilli 32 intervalles temporels inter-battements par bloc et participant (cf. 3.1.1, p. 87).

Les données de la tâche de TMS ont fourni deux mesures représentatives de la capacité des participants à produire une séquence rythmique :

1. Le TMS moyen, représentatif de la période de référence
2. La stabilité du TMS, représentative des capacités de traitement du rythme

Nous n'avons pas cherché à observer l'effet de variables indépendantes sur le TMS moyen ou la stabilité du TMS, seulement la production spontanée des participants. Nous présentons ainsi la distribution et la tendance centrale générale de ces mesures. De plus, à la suite de Drake et al. (2000) qui rapportent une incidence de l'âge ou du niveau d'expertise musical sur le TMS, nous avons pensé intéressant d'explorer les liens possibles entre nos deux mesures et les caractéristiques socio-démographiques des participants (données du questionnaire, cf. 6.1, p. 113). Nous avons donc intégré les observations concernant l'âge, le sexe, la pratique musicale, de la danse et des jeux vidéo (Oui/Non) et le nombre d'années de pratique de la danse et de la musique dans les analyses descriptives et inférentielles. Nous présentons les résultats (statistiques descriptives puis inférentielles) du TMS moyen en partie 8.1.1 et de la stabilité du TMS (statistiques descriptives puis inférentielles) en partie 8.1.2 (p. 139).

8.1.1 TEMPO MOTEUR SPONTANE MOYEN

Le Tempo Moteur Spontané (TMS) moyen nous renseigne sur la période de référence, c'est-à-dire la vitesse de production rythmique la plus naturelle pour le participant. Plus la valeur du TMS moyen est petite (par exemple, 500 ms⁷⁴), plus la période de référence du

⁷⁴ Un TMS moyen de 500 ms = 2 Hz, soit une vitesse de deux battements de la main par seconde.

participant est rapide. Au contraire, une valeur du TMS moyen élevée (par exemple 1.200 ms⁷⁵) indique un TMS moyen plus lent.

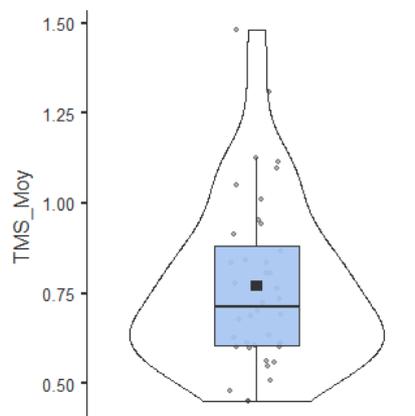
8.1.1.1 EXPLORATION DESCRIPTIVE

Le TMS moyen est donné par la moyenne des intervalles temporels inter-battements (IBI⁷⁶) de chaque participant. Le recueil des IBI des participants s'étant effectué sur deux blocs expérimentaux de 32 IBI⁷⁷ chacun (cf. 3.1.1, p. 87), nous avons calculé pour chaque participant la moyenne des IBI de chaque bloc, ensuite agrégées en une seule valeur moyenne, le TMS moyen du participant. A partir de l'ensemble des TMS moyens des individus ($n = 40$), nous avons pu calculer le TMS moyen global de notre échantillon. La distribution des TMS moyens des individus ne suivant pas une distribution normale ($p = 0,009$ au test de normalité de Shapiro-Wilk, cf. **Figure 30** et **Tableau 8**), nous avons privilégié la médiane par rapport à la moyenne comme indice de tendance centrale.

Le TMS moyen global pour l'ensemble des participants est donc de 712 ms.

Figure 30 :

Représentation de la distribution des TMS moyens des participants



Note. $N = 40$. La boîte à moustaches indique la médiane et les quartiles 1 et 3 ; le carré la moyenne ; le violon la densité des TMS moyens (en secondes).

⁷⁵ Un TMS de 1.200 ms = 0,83 Hz, soit une vitesse de 0,83 battement par seconde.

⁷⁶ De l'anglais *Inter-Beat Interval*.

⁷⁷ Auxquels ont été soustraites les valeurs aberrantes (cf. 6.2.2, p. 34).

Tableau 8 :
Valeurs descriptives des TMS moyens des participants

Statistiques descriptives

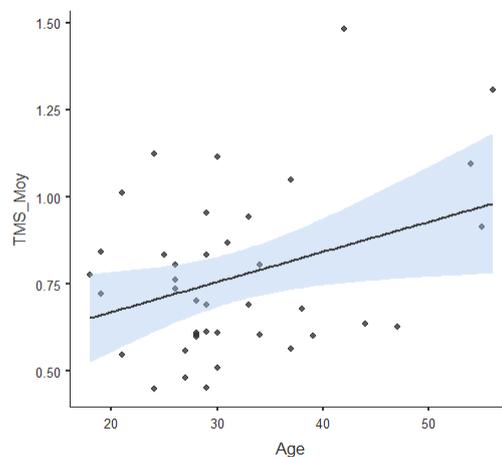
| | N | Manquants | Moyenne | Médiane | Ecart-type | Minimum | Maximum | Shapiro-Wilk | |
|---------|----|-----------|---------|---------|------------|---------|---------|--------------|-------|
| | | | | | | | | W | p |
| TMS_Moy | 40 | 0 | 0.770 | 0.712 | 0.234 | 0.450 | 1.48 | 0.922 | 0.009 |

Note. Une valeur $p < 0,05$ au test de Shapiro-Wilk indique une distribution non paramétrique des données, invitant à considérer de préférence la médiane par rapport à la moyenne (TMS_moy, en secondes).

Pour rappel, nous avons également souhaité observer l'effet des variables socio-démographiques sur le TMS moyen, et avons ainsi intégré dans nos analyses les observations concernant l'effet de l'âge, du sexe, de la pratique musicale, de la danse et des jeux vidéo (Oui/Non) et du nombre d'années de pratique de la danse et de la musique des participants⁷⁸. Notre exploration descriptive semble ainsi indiquer plusieurs tendances.

Plus l'âge des participants est élevé, plus leur TMS moyen est lent (cf. **Figure 31**).

Figure 31 :
Evolution du TMS moyen en fonction de l'âge.



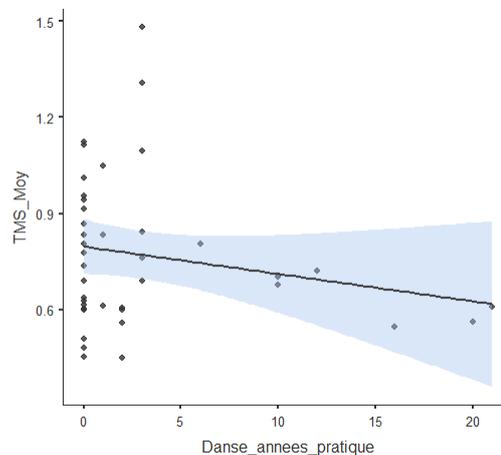
Note. La ligne indique la droite de régression linéaire et l'espace coloré l'erreur standard (TMS_moy, en secondes).

⁷⁸ Pour rappel, dans cette étude, nous employons les termes « musicien », « danseur » ou « joueur de jeux vidéo » dès lors qu'un participant a indiqué pratiquer ou avoir pratiqué une activité musicale, de danse ou de jeux vidéo respectivement.

Plus le nombre d'années de pratique de la danse des participants est élevé, plus leur TMS moyen est rapide (cf. **Figure 32**).

Figure 32 :

Evolution du TMS moyen en fonction du nombre d'années de pratique de la danse



Note. La ligne indique la droite de régression linéaire et l'espace coloré l'erreur standard (TMS_moy, en secondes).

De plus, il semblerait que les femmes présentent un TMS moyen plus rapide (médiane TMS = 696 ms) que les hommes (médiane TMS = 785 ms) ; que les participants musiciens présentent un TMS moyen plus rapide (médiane TMS = 702 ms) que les non-musiciens (médiane TMS = 777 ms) ; que les participants musiciens percussionnistes présentent un TMS moyen le plus rapide (médiane TMS = 525 ms) et les guitaristes le TMS moyen le plus lent (médiane TMS = 868 ms) ; que les participants danseurs présentent un TMS moyen plus rapide (médiane TMS = 696 ms) que les non-danseurs (médiane TMS = 756 ms) ; et que les participants joueurs de jeux vidéo présentent un TMS moyen plus rapide (médiane TMS = 712 ms) que les non-joueurs (médiane TMS = 766 ms).

8.1.1.2 ANALYSE INFERENTIELLE

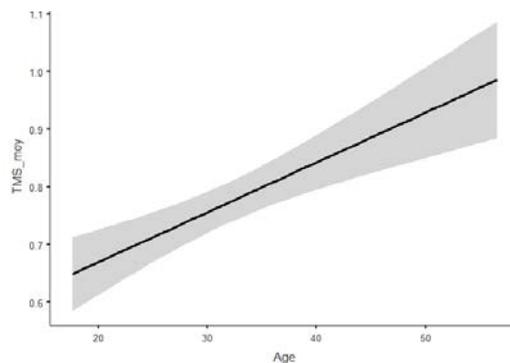
Afin de modéliser la relation entre les caractéristiques socio-démographiques et le TMS moyen des participants, nous avons appliqué un modèle linéaire mixte à nos données ($n = 80^{79}$), avec en effets fixes les variables socio-démographiques, en effet aléatoire les participants et en variable dépendante le TMS moyen. Pour sélectionner le modèle linéaire mixte estimant au mieux les effets des paramètres socio-démographiques sur les variations du TMS moyen, nous

⁷⁹ 1 valeur de TMS moyen * 2 blocs * 40 participants.

avons suivi la procédure en pas à pas descendants (« *top-down* ») préconisée par Zuur et al. (2009) (nous avons détaillé cette procédure en partie 7.1, p. 124). Nous sommes ainsi partie d'un modèle saturé avec les paramètres suivants (effets fixes) : covariables âge, années de pratique de la musique et de la danse ; facteurs sexe, pratique de la musique, de la danse et des jeux vidéo. A partir de ce modèle saturé en effets fixes et après avoir sélectionné la structure aléatoire la plus adaptée, nous avons enlevé un à un les paramètres fixes pour arriver au modèle le plus simple et le plus ajusté. La comparaison détaillée des modèles jusqu'à parvenir au modèle le plus ajusté est présentée en ANNEXES – PARTIE 2 (document complémentaire, 1.1.1.1, p. 17).

Nous observons un effet significatif de l'âge sur le TMS moyen des participants ($Estimate = 0,009$; $p < 0,029$), qui confirme la tendance que nous avons observée lors de l'exploration descriptive. Les participants les plus âgés sont ainsi ceux qui ont un TMS moyen le plus lent (cf. **Figure 33**).

Figure 33 :
Evolution du TMS moyen en fonction de l'âge



Note. . La droite représente l'estimation (coefficient de régression estimé = 0,009) de la modélisation, la zone grisée l'erreur standard (TMS_moy, en secondes).

8.1.1.3 SYNTHÈSE DU TMS MOYEN

L'exploration descriptive du TMS moyen a montré une tendance centrale pour le TMS moyen global (médiane des TMS moyens individuels) de 712 ms ainsi que plusieurs tendances. Le TMS moyen tend à ralentir avec l'avancée de l'âge et à accélérer avec l'augmentation du nombre d'années de pratique de la danse. Le TMS moyen tend à être plus rapide chez les femmes, les musiciens (et notamment les percussionnistes), les danseurs et les joueurs de jeux vidéo.

D'après les résultats du modèle inférentiel appliqué à nos données, nous pouvons nuancer les tendances initialement observées : seul l'âge présente un effet significatif sur la

variabilité du TMS moyen, les participants les plus âgés ayant un TMS moyen plus lent. Cependant, en raison du caractère non contrôlé des variables socio-démographiques dans notre étude et du nombre restreint d'observations, l'effet simple de l'âge est néanmoins à interpréter avec précaution.

Nous retrouvons ainsi une valeur de la période de référence proche de celles rapportées dans la littérature (400 à 800 ms) (cf. 1.2.1.3.1, p. 19) ainsi qu'un effet de l'âge, le TMS moyen étant plus lent chez les participants plus âgés (tout comme Dalla Bella et al., 2017 ; Drake et al., 2000).

Ces résultats concernant le TMS moyen seront ensuite mis en regard des observations concernant la tâche de Synchronisation Sensori-Motrice (SMS) (Activité 1) et l'Activité 2 afin d'investiguer l'*Hypothèse 1*, à propos de l'existence d'une période de référence supramodale. En effet, nous nous attendons à ce que les tempi les plus stables et précis dans la tâche de SMS (parmi les trois tempi proposés : 225, 600 et 900 ms) se rapprochent du TMS moyen. De plus, nous nous attendons également à ce que la valeur du TMS moyen soit corrélée avec la valeur du Temps d'Initiation (TI) de la parole moyen, que nous mesurons dans l'Activité 2. Nous reviendrons sur ces comparaisons entre tâches dans la partie 10 (p. 187).

8.1.2 STABILITE DU TEMPO MOTEUR SPONTANE

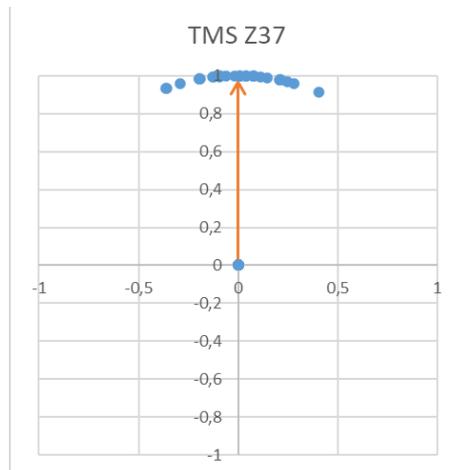
Tout comme le TMS moyen, la stabilité du TMS a été mesurée à partir des Intervalles Inter-Battements (IBI) du TMS de chaque participant. La caractérisation de la stabilité s'est faite à l'aide des statistiques circulaires (cf. 6.2.2, p. 115) : le TMS moyen de chacun des participants a été converti en une échelle circulaire (TMS moyen du participant en ms = 360°) de rayon 1 sur laquelle les IBI ont été répartis selon leur asynchronie par rapport au TMS moyen (cf. par exemple **Figure 34**). Ainsi, un IBI correspondant exactement au TMS moyen représente une asynchronie nulle (0°) et est placé pile au « nord » du cercle, tandis qu'un intervalle correspondant à un contre-temps par rapport au TMS moyen sera placé à l'opposé sur le cercle (à 180° , soit au « sud » du cercle). La stabilité moyenne du participant peut ensuite être calculée à partir de l'inverse de la variance circulaire des asynchronies, résultant en un vecteur de longueur moyenne \bar{R} comprise entre 0 (instable) et 1 (stable)⁸⁰.

⁸⁰ 1 correspondant à la taille du rayon du cercle.

La tâche de TMS s'étant déroulée sur deux blocs expérimentaux (cf. 3.1.1, p. 87), nous avons ainsi recueilli deux valeurs par participant pour la mesure de la stabilité du TMS.

Figure 34 :

Représentation de la répartition des IBI et du vecteur du participant Z37



Note. Représentation de la répartition des IBI (points bleus en arc de cercle) et du vecteur (flèche orange) du participant Z37.

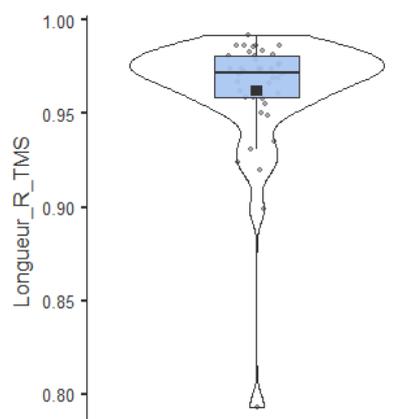
8.1.2.1 EXPLORATION DESCRIPTIVE

La longueur moyenne \bar{R} , qui représente la stabilité du participant, a été calculée pour chacun des deux blocs expérimentaux pour chaque participant (cf. 3.1.1, p. 87). Afin de rejeter l'hypothèse nulle d'une répartition aléatoire (uniforme) des IBI du participant, nous avons appliqué le test de Rayleigh à la longueur moyenne \bar{R} de chaque bloc pour chaque participant. Il en est ressorti que tous les participants ont une distribution de leurs frappes significativement différente d'une distribution aléatoire ($p < 0,001$ pour les deux blocs expérimentaux) et sont donc capables de produire une séquence rythmique régulière. Nous avons agrégé les deux valeurs de stabilité moyenne de chaque participant (une par bloc expérimental) afin d'en faire la moyenne, représentative de la stabilité moyenne du participant à la tâche de TMS, tous blocs confondus.

A partir de l'ensemble des valeurs agrégées de stabilité moyenne de chaque participant ($n = 40$), nous avons pu calculer la stabilité moyenne globale, médiane des stabilités moyennes individuelles de notre échantillon. La distribution des stabilités moyennes des individus ne suivant pas une distribution normale ($p < 0,001$ au test de normalité de Shapiro-Wilk, cf. **Figure 35** et **Tableau 9**), nous avons privilégié la médiane par rapport à la moyenne comme indice de tendance centrale.

La stabilité globale du TMS pour l'ensemble de notre échantillon est donc de $\bar{R} = 0,971$.

Figure 35 :
Représentation graphique de la stabilité du TMS des participants



Note. Représentation graphique de la stabilité du TMS des participants (Longueur_R_TMS). La boîte à moustaches indique la médiane et les quartiles 1 et 3 ; le carré la moyenne ; le violon la densité.

Tableau 9 :
Valeurs descriptives des stabilités moyennes du TMS

Statistiques descriptives

| | N | Moyenne | Médiane | Ecart-type | Minimum | Maximum | Shapiro-Wilk | |
|----------------|----|---------|---------|------------|---------|---------|--------------|--------|
| | | | | | | | W | p |
| Longueur_R_TMS | 40 | 0.962 | 0.972 | 0.0341 | 0.793 | 0.992 | 0.648 | < .001 |

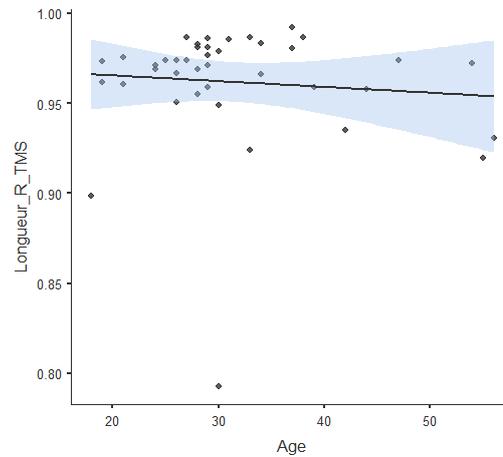
Note. Un p inférieur à 0,05 au test de Shapiro-Wilk indique une distribution non paramétrique des données. Longueur_R_TMS : stabilité moyenne du TMS.

A la suite de nos observations sur le TMS moyen, nous avons également souhaité observer l'effet des variables socio-démographiques sur la stabilité du TMS, et avons ainsi intégré dans nos analyses les observations concernant l'effet de l'âge, du sexe, de la pratique musicale, de la danse et des jeux vidéo (Oui/Non) et du nombre d'années de pratique de la danse et de la musique des participants⁸¹. Notre exploration descriptive semble ainsi indiquer plusieurs tendances.

Plus l'âge des participants est élevé, moins leur TMS est stable (cf. **Figure 36**).

⁸¹ Pour rappel, dans cette étude, nous employons les termes « musicien », « danseur » ou « joueur de jeux vidéo » dès lors qu'un participant a indiqué pratiquer ou avoir pratiqué une activité musicale, de danse ou de jeux vidéo respectivement.

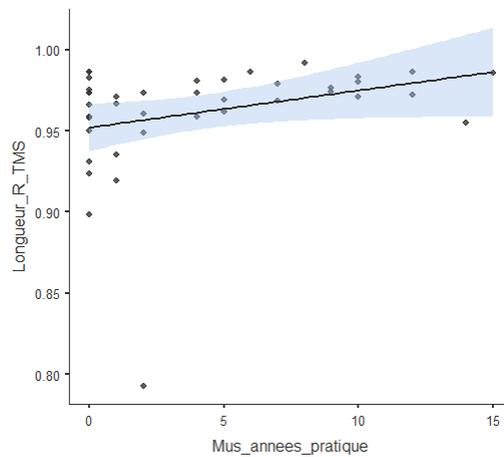
Figure 36 :
Evolution de la stabilité du TMS en fonction de l'âge



Note. La ligne indique la droite de régression linéaire et l'espace coloré l'erreur standard. Longueur_R_TMS : stabilité moyenne du TMS.

Plus le nombre d'années de pratique de la musique ou de la danse des participants est élevé, plus leur TMS est stable (cf. **Figure 37** et **Figure 38**).

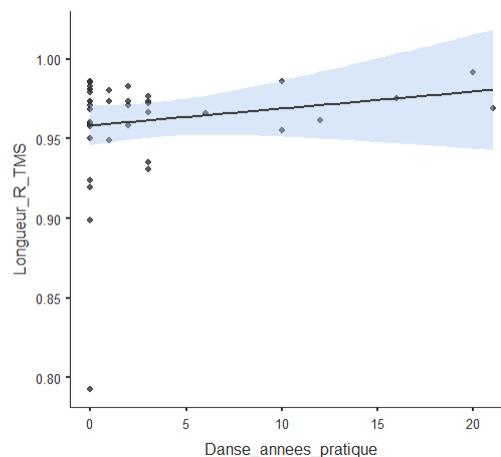
Figure 37 :
Evolution de la stabilité du TMS en fonction du nombre d'années de musique



Note. La ligne indique la droite de régression linéaire et l'espace coloré l'erreur standard. Longueur_R_TMS : stabilité moyenne du TMS.

Figure 38 :

Evolution de la stabilité du TMS selon le nombre d'années de pratique de la danse



Note. La ligne indique la droite de régression linéaire et l'espace coloré l'erreur standard. Longueur_R_TMS : stabilité moyenne du TMS.

De plus, il semblerait que les participants percussionnistes soient les plus stables tous tempi confondus (médiane = 0,976), suivis de près par les guitaristes et les participants pratiquant plusieurs instruments, tandis que les flûtistes seraient les moins stables (médiane = 0,959).

8.1.2.2 ANALYSE INFERENTIELLE

Afin de modéliser la relation entre les caractéristiques socio-démographiques et la stabilité moyenne du TMS des participants, nous avons appliqué un modèle linéaire mixte à nos données de stabilité ($n = 80^{82}$), avec en effets fixes les variables socio-démographiques, en effet aléatoire les participants et en variable dépendante la stabilité moyenne du TMS. Pour sélectionner le modèle estimant au mieux les effets des paramètres socio-démographiques sur les variations de stabilité du TMS, nous avons procédé de la même manière que pour les variations du TMS moyen (cf. 8.1.1, p. 134). Nous sommes partie d'un modèle saturé avec les paramètres suivants (effets fixes) : covariables âge, années de pratique de la musique et de la danse ; facteurs sexe, pratique de la musique, de la danse et des jeux vidéo. Après avoir comparé pas à pas la pertinence des paramètres pour parvenir à un modèle le plus simple et le plus ajusté (cf. ANNEXES – PARTIE 2 document complémentaire, 1.1.1.2, p. 26), nous avons appliqué le nouveau modèle à nos données.

⁸² 40 participants * 2 blocs

La modélisation de la relation entre les paramètres socio-démographiques et la stabilité moyenne du TMS des participants n'a pas permis d'éclairer les tendances exposées dans l'exploration descriptive, aucun paramètre fixe n'ayant d'effet significatif sur la variabilité de la stabilité moyenne du TMS.

8.1.2.3 SYNTHÈSE DE LA STABILITÉ DU TMS

L'exploration descriptive de la stabilité moyenne du TMS a montré une tendance centrale pour la stabilité du TMS globale (médiane) de 0,971⁸³ ainsi que plusieurs tendances. La stabilité moyenne du TMS tend à diminuer avec l'âge mais à augmenter avec l'augmentation du nombre d'années de pratique de la musique et de la danse. La stabilité du TMS tend à être la plus élevée chez les percussionnistes.

Les résultats du modèle inférentiel n'ont pas montré d'effet significatif des paramètres socio-démographiques sur la stabilité du TMS. En raison du caractère non contrôlé des variables socio-démographiques dans notre étude et du nombre restreint d'observations, l'absence d'effets simples est néanmoins à interpréter avec précaution.

Nous retrouvons ainsi les valeurs rapportées dans la littérature quant à la stabilité de la pulsation. Canette et al. (2020) rapportent ainsi une stabilité moyenne de la synchronisation sensori-motrice aux alentours de 0,85. Sowiński et Dalla Bella (2013) rapportent dans les résultats de leur première étude une valeur de stabilité moyenne (synchronisation à un rythme de métronome) de 0,94 et de 0,83 (synchronisation à un rythme musical).

Ces résultats concernant la stabilité du TMS seront ensuite mis en regard des observations concernant la tâche de Synchronisation Sensori-Motrice (SMS) (Activité 1) et l'Activité 2 afin d'investiguer l'Hypothèse 6, à propos de l'existence de mécanismes cognitifs partagés pour le traitement du rythme et du langage. En effet, nous nous attendons à ce que les participants les plus stables et précis dans les tâches de TMS et SMS soient ceux qui présentent le moins d'erreurs de lecture et un Temps d'Initiation (TI) de la parole plus rapide. Nous reviendrons sur ces comparaisons entre tâches dans la partie 10 (p. 187).

⁸³ Pour rappel, la stabilité du participant est donnée par un vecteur de longueur moyenne \bar{R} , comprise entre 0 (instable) et 1 (stable).

8.2 TACHE DE SYNCHRONISATION SENSORI-MOTRICE

Dans la tâche de Synchronisation Sensori-Motrice (SMS), les participants avaient pour consigne de battre la pulsation en appuyant avec leur doigt sur une touche, en rythme avec un métronome. La mesure de la SMS s'effectuait selon des séquences de métronome à trois tempi (225, 600 et 900 ms) de 40,5 secondes chacune, chaque tempo étant présenté deux fois, dans un ordre aléatoirisé entre les participants (cf. 3.1.2, p. 87).

La tâche de Synchronisation Sensori-Motrice (SMS) nous a permis de recueillir deux mesures représentatives de la capacité des participants à percevoir une séquence rythmique et à s'y synchroniser :

1. La stabilité de la SMS
2. La précision de la SMS

Ces deux mesures ont été investies dans trois modalités de tempo, soit des séquences rythmiques d'ISI⁸⁴ 225 ms, 600 ms et 900 ms.

Nous avons cherché à observer l'effet du tempo (variable indépendante) sur la stabilité et la précision de la SMS, afin d'investiguer l'*Hypothèse 1.1*. En effet, nous nous attendons à ce que le tempo auquel la SMS d'un participant est la plus stable et/ou précise se rapproche du Tempo Moteur Spontané (TMS) moyen.

De plus, les mesures de stabilité et précision de la SMS renseignent sur les capacités de traitement du rythme, que nous mettons en regard des capacités de traitement du langage dans la partie 10 (p. 187), afin d'investiguer l'*Hypothèse 6*. En effet, nous supposons que les participants ayant de meilleures capacités de traitement du rythme sont également ceux qui présentent les meilleures capacités de traitement du langage.

Tout comme pour le TMS, nous proposons à la suite de Drake et al. (2000) ou Repp (2005), qui rapportent une incidence de l'âge ou du niveau d'expertise musical sur la SMS, d'explorer également les liens possibles entre les mesures de stabilité et précision de la SMS et les caractéristiques socio-démographiques des participants (données du questionnaire, cf. 6.1, p. 111). Nous avons donc intégré les observations concernant l'âge, le sexe, la pratique

⁸⁴ De l'anglais *Inter-Stimulus Interval*.

musicale, de la danse et des jeux vidéo (Oui/Non) et le nombre d'années de pratique de la danse et de la musique dans les analyses descriptives et inférentielles.

Nous présentons les résultats (statistiques descriptives et inférentielles) de la stabilité de SMS en partie 8.2.1 et de la précision de la SMS en partie 8.2.2 (p. 156).

8.2.1 STABILITE DE LA SYNCHRONISATION SENSORI-MOTRICE

Tout comme la stabilité du TMS (cf. 8.1.2, p. 139), la stabilité de la SMS est mesurée à partir des statistiques circulaires et plus précisément à partir de la longueur résultante moyenne \bar{R} du vecteur circulaire. Les participants étaient exposés à chaque tempo sur deux blocs expérimentaux (cf. 3.1.2, p. 87), nous avons ainsi recueilli deux valeurs de stabilité moyenne par participant et par tempo pour la mesure de la stabilité de la SMS.

8.2.1.1 EXPLORATION DESCRIPTIVE

Après avoir calculé les longueurs moyennes \bar{R} , représentatives de la stabilité, pour chaque participant ($n = 40$), chaque tempo ($n = 3$) et chaque bloc expérimental ($n = 2$), soit un total de 240 longueurs \bar{R} , nous avons vérifié leur significativité avec le test de Rayleigh⁸⁵. Il en est ressorti que 94,6 % des longueurs \bar{R} des participants sont significatives (cf. **Tableau 10**). Nous avons donc écarté les 5,4 % non significatives pour ne pas biaiser nos calculs ultérieurs de stabilité et précision de la SMS. Il est également intéressant de noter que la quasi-totalité des longueurs \bar{R} non significatives sont associées au tempo le plus rapide (225 ms) : sur les 13 longueurs non significatives au test de Rayleigh, 12 sont associées au tempo 225 ms (cf. 'n.s.' dans le **Tableau 10**), indiquant une plus grande difficulté à mettre en place la synchronisation au tempo de 225 ms.

⁸⁵ Pour rappel, le test de Rayleigh permet de distinguer la répartition des frappes du participant d'une répartition aléatoire, et témoigne de la réussite de la mise en place de la synchronisation.

Tableau 10 :*Tableau de fréquences de la significativité des longueurs \bar{R} de la SMS*

Fréquences de Significativité

| Significativité | Tempo | Quantités | % du Total | % cumulés |
|-----------------|-------|-----------|------------|-----------|
| ** | 225ms | 5 | 2.1 % | 2.1 % |
| | 600ms | 3 | 1.3 % | 3.3 % |
| | 900ms | 0 | 0.0 % | 3.3 % |
| *** | 225ms | 63 | 26.3 % | 29.6 % |
| | 600ms | 76 | 31.7 % | 61.3 % |
| | 900ms | 80 | 33.3 % | 94.6 % |
| n.s. | 225ms | 12 | 5.0 % | 99.6 % |
| | 600ms | 1 | 0.4 % | 100.0 % |
| | 900ms | 0 | 0.0 % | 100.0 % |

Note. La significativité des longueurs \bar{R} est calculée à partir du test de Rayleigh et appliquée aux trois tempi de métronome (225, 600 et 900 ms). ‘***’ indique un $p \leq 0,01$, ‘****’ $p \leq 0,001$ et ‘n.s.’ $p > 0,05$.

Les longueurs \bar{R} pour chaque tempo ont été calculées à partir de la moyenne des deux blocs expérimentaux de chaque participant (soit $n = 40$ pour chaque tempo). En cas de valeur manquante (longueur \bar{R} non significative au test de Rayleigh) pour un des deux blocs expérimentaux d’un tempo, nous avons utilisé la valeur restante comme moyenne de ce tempo pour le participant. En cas de valeur manquante pour chacun des deux blocs d’un tempo, nous n’avons pas estimé la stabilité de ce tempo pour le participant. Nous avons ensuite agrégé les moyennes des trois tempi pour calculer la stabilité moyenne des participants tous tempi confondus. Lorsqu’un participant avait une valeur manquante pour l’un de ses tempi⁸⁶, nous n’avons pas pu calculer de stabilité moyenne tous tempi confondus pour le participant.

A partir de l’ensemble des stabilités moyennes de chaque participant tous tempi confondus ($n = 38^{87}$), nous avons pu calculer la stabilité globale, médiane des stabilités moyennes individuelles de notre échantillon. La distribution des stabilités moyennes agrégées des individus ne suivant pas une distribution normale ($p < 0,001$ au test de normalité de Shapiro-Wilk, cf. **Figure 39** et **Tableau II**), nous avons privilégié la médiane par rapport à la moyenne comme indice de tendance centrale.

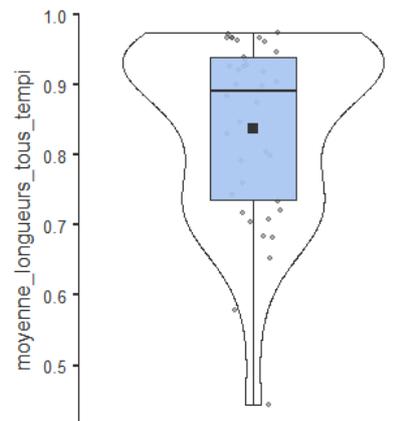
⁸⁶ Une valeur manquante à un tempo résulte de deux longueurs non significatives au test de Rayleigh pour ce tempo.

⁸⁷ 40 participants * 1 observation, sans les valeurs manquantes.

La stabilité moyenne globale de la SMS pour l'ensemble de notre échantillon tous tempi confondus est de $\bar{R} = 0,891$.

Figure 39 :

Représentation graphique de la stabilité moyenne de la SMS des participants



Note. La boîte à moustaches indique la médiane et les quartiles 1 et 3 ; le carré la moyenne ; le violon la densité. moyenne_longueur_tous_tempi : stabilité moyenne de la SMS, tous tempi confondus.

Tableau 11 :

Valeurs descriptives de la stabilité moyenne de la SMS

Statistiques descriptives

| | N | Manquants | Moyenne | Médiane | Ecart-type | Minimum | Maximum | Shapiro-Wilk | |
|------------------------------|----|-----------|---------|---------|------------|---------|---------|--------------|--------|
| | | | | | | | | W | p |
| moyenne_longueurs_tous_tempi | 38 | 2 | 0.838 | 0.891 | 0.129 | 0.443 | 0.975 | 0.881 | < .001 |

Note. Valeurs descriptives des longueurs résultantes \bar{R} , représentatives de la stabilité moyenne de la SMS, tous tempi confondus. Une valeur de $p < 0,05$ au test de Shapiro-Wilk indique une distribution non paramétrique des données. moyenne_longueur_tous_tempi : stabilité moyenne de la SMS, tous tempi confondus.

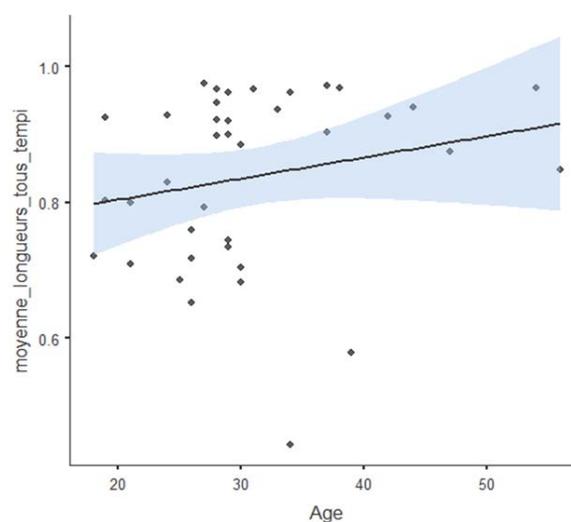
De plus, en explorant les données, nous observons plusieurs tendances concernant la stabilité de la SMS, selon le tempo du métronome (ISI 225, 600 ou 900 ms). Le tempo auquel les participants sont les plus stables est 900 ms (médiane = 0,961), suivi par le tempo 600 ms (médiane = 0,941) puis le tempo 225 ms (médiane = 0,769).

Pour rappel, nous avons également souhaité observer l'effet des variables socio-démographiques sur la stabilité de la SMS, et avons ainsi intégré dans nos analyses les observations concernant l'effet de l'âge, du sexe, de la pratique musicale, de la danse et des jeux vidéo (Oui/Non) et du nombre d'années de pratique de la danse et de la musique des participants⁸⁸. Notre exploration descriptive semble ainsi indiquer plusieurs tendances.

La stabilité de la SMS augmente avec l'âge et le nombre d'années de pratique de la musique et de la danse (cf. **Figure 40**, **Figure 41** et **Figure 42**)

Figure 40 :

Evolution de la stabilité de la SMS en fonction de l'âge

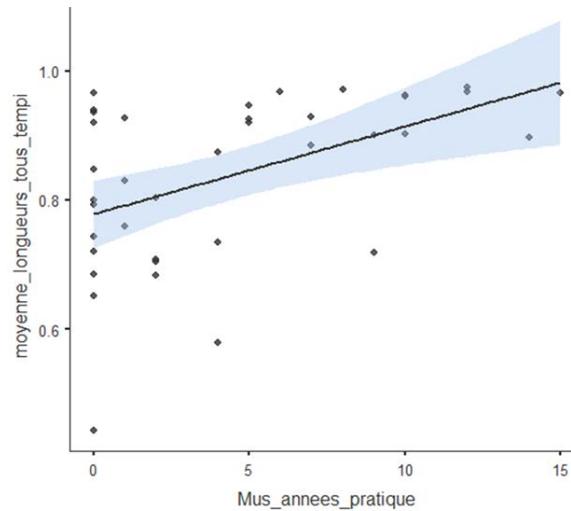


Note. . La ligne indique la droite de régression linéaire et l'espace coloré l'erreur standard. moyenne_longueur_tous_tempi : stabilité moyenne de la SMS, tous tempi confondus.

⁸⁸ Pour rappel, dans cette étude, nous employons les termes « musicien », « danseur » ou « joueur de jeux vidéo » dès lors qu'un participant a indiqué pratiquer ou avoir pratiqué une activité musicale, de danse ou de jeux vidéo respectivement.

Figure 41 :

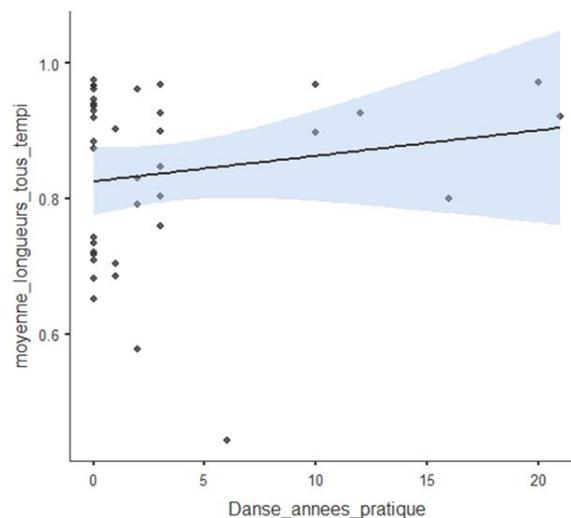
Evolution de la stabilité de la SMS en fonction du nombre d'années de musique



Note. Evolution de la stabilité de la SMS en fonction du nombre d'années de pratique de la musique. La ligne indique la droite de régression linéaire et l'espace coloré l'erreur standard. `moyenne_longueurs_tous_tempi` : stabilité moyenne de la SMS, tous tempi confondus.

Figure 42 :

Evolution de la stabilité de la SMS en fonction du nombre d'années de pratique de la danse



Note. Evolution de la stabilité de la SMS en fonction du nombre d'années de pratique de la danse. La ligne indique la droite de régression linéaire et l'espace coloré l'erreur standard. `moyenne_longueurs_tous_tempi` : stabilité moyenne de la SMS, tous tempi confondus.

De plus, il semblerait que la stabilité de la SMS soit plus élevée chez les hommes (médiane = 0,932) que les femmes (médiane = 0,803) ; que la stabilité de la SMS soit plus élevée chez les musiciens (médiane = 0,901) que les non-musiciens (médiane = 0,792) ; et que

la stabilité de la SMS soit la plus élevée chez les participants musiciens percussionnistes (médiane = 0,955), tandis que les guitaristes sont les moins stables (médiane = 0,734).

8.2.1.2 ANALYSE INFERENTIELLE

Afin d'investiguer l'*Hypothèse 1* et de modéliser la relation entre le tempo et la stabilité de la SMS, nous avons ensuite appliqué un modèle linéaire mixte à nos données de stabilité ($n = 227^{89}$).

Dans un premier temps, nous avons appliqué à nos données un modèle comprenant uniquement notre variable indépendante d'intérêt (le tempo) (cf. 8.2.1.2.1). Dans un second temps, nous avons construit et appliqué un second modèle incluant également des variables contrôles socio-démographiques afin d'estimer la potentielle interaction entre les variables socio-démographiques et notre effet d'intérêt (le tempo) (cf. 8.2.1.2.2, p. 152).

8.2.1.2.1 EFFET DU TEMPO SUR LA STABILITE DE LA SMS

Afin de modéliser la relation entre le tempo et la stabilité moyenne de la SMS des participants, nous avons appliqué un modèle linéaire mixte à nos données de stabilité ($n = 227$), avec en effets fixes le tempo, en effet aléatoire les participants et en variable dépendante la stabilité de la SMS.

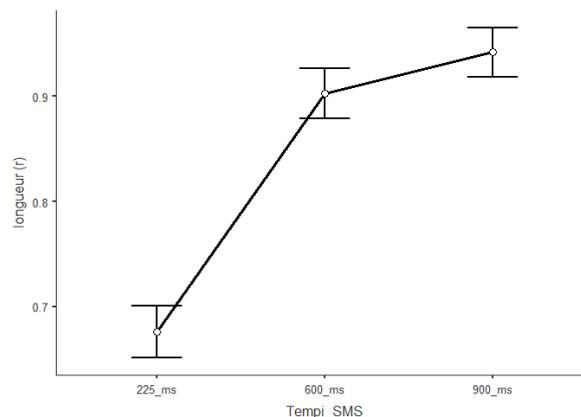
Après avoir établi la structure aléatoire la plus ajustée à notre modèle (cf. ANNEXES – PARTIE 2, document complémentaire, 1.1.2.1, p. 35), nous avons appliqué celui-ci à nos données de stabilité de la SMS.

Nous observons un effet principal du tempo ($p < 0,001$) sur la stabilité de la SMS. Les comparaisons post-hoc (correction de Bonferroni) montrent une différence significative entre d'une part le tempo 225 ms et d'autre part les tempi 600 ms (*Difference* = -0,227 ; $p < 0,001$) et 900 ms (*Difference* = -0,266 ; $p < 0,001$), la stabilité de la SMS étant moins élevée pour le tempo 225 ms que pour les tempi 600 et 900 ms (cf. **Figure 43**). La différence de stabilité entre les tempi 600 ms et 900 ms n'est pas significative ($p = 0,259$).

⁸⁹ Données réelles/désagrégées : 40 participants * 3 tempi * 2 blocs, sans les valeurs manquantes.

Figure 43 :

Comparaison de la stabilité de la SMS en fonction du tempo (1^{er} modèle)



Note. Les droites représentent l'estimation (différence des moyennes estimées) de la modélisation et les barres l'erreur standard. longueur(r) : stabilité moyenne de la SMS.

8.2.1.2.2 EFFET DU TEMPO ET DES CARACTERISTIQUES SOCIO-DEMOGRAPHIQUES SUR LA STABILITE DE LA SMS

Après avoir estimé l'effet du tempo (variable d'intérêt) sur la stabilité de la SMS, nous avons voulu aller plus loin en ajoutant au modèle des caractéristiques socio-démographiques (variables contrôles) qui permettraient de contrôler que la variance observée est bien liée aux variations de tempo.

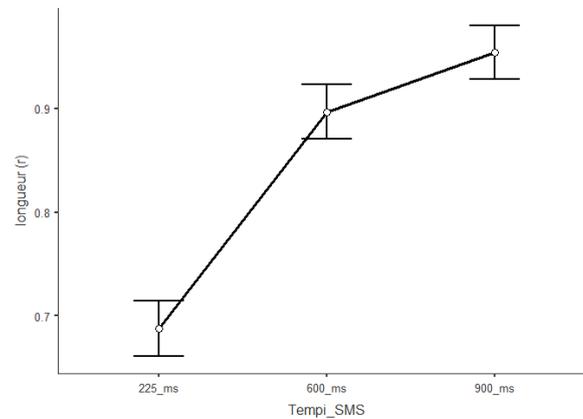
Nous sommes donc partie d'un modèle saturé avec les paramètres suivants : facteur d'intérêt tempo ; covariables contrôles âge, années de pratique de la musique et de la danse ; facteurs contrôles sexe, pratique de la musique, de la danse et des jeux vidéo ; et les interactions doubles entre d'une part le facteur d'intérêt tempo et d'autre part les variables contrôles. Après avoir comparé pas à pas la pertinence des paramètres pour parvenir à un modèle le plus simple et le plus ajusté (cf. ANNEXES – PARTIE 2, document complémentaire, 1.1.2.1, p. 38), nous avons appliqué le nouveau modèle à nos données.

Nous observons à nouveau un effet principal du tempo ($p < 0,001$) sur la stabilité de la SMS, qui est modulé par l'âge ($p = 0,016$ pour l'interaction « Temps_SMS : Age ») et le nombre d'années de pratique de la musique ($p < 0,001$ pour l'interaction « Temps_SMS : Mus_annees_pratique »). Les comparaisons post-hoc (correction de Bonferroni) montrent une différence significative entre d'une part le tempo 225 ms et d'autre part les tempi 600 ms ($Difference = -0,209$; $p < 0,001$) et 900 ms ($Difference = -0,267$; $p < 0,001$), ainsi qu'entre les tempi 600 ms et 900 ms ($Difference = -0,058$; $p = 0,039$). La stabilité de la SMS est ainsi moins

élevée pour le tempo 225 ms que pour les tempi 600 et 900 ms, et moins élevée pour le tempo 600 ms que le tempo 900 ms (cf. **Figure 44**).

Figure 44 :

Comparaison de la stabilité de la SMS en fonction du tempo (2^{ème} modèle)



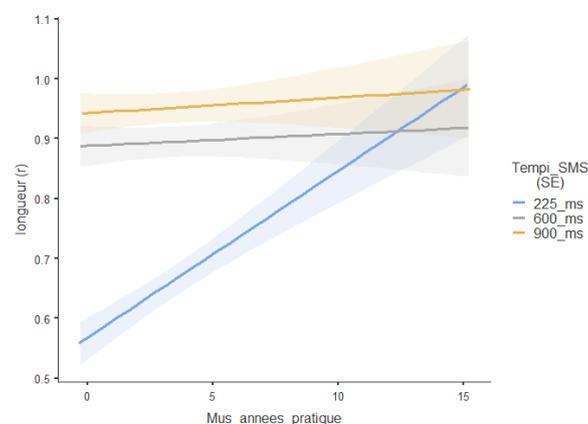
Note. Les droites représentent l'estimation (différence des moyennes estimées) de la modélisation et les barres l'erreur standard. longueur(r) : stabilité moyenne de la SMS.

Les différences de stabilité de la SMS en fonction du tempo sont modulées par l'âge ($p = 0,016$). Cependant, nous n'avions pas posé d'hypothèses spécifiques au préalable quant aux niveaux d'âge impactés et ne regardons ainsi pas plus en détail cette interaction.

La stabilité de la SMS augmente avec l'augmentation du nombre d'années de pratique de la musique pour le tempo 225 ms ($Estimate = 0,028$; $p < 0,001$), tandis que l'effet du nombre d'années de pratique de la musique n'est pas significatif pour les tempi 600 ms ($p = 0,760$) et 900 ms ($p = 0,690$) (cf. **Figure 45**).

Figure 45 :

Evolution de la stabilité de la SMS selon les tempi et le nombre d'années de musique



Note. Evolution de la stabilité de la SMS en fonction du nombre d'années de pratique de la musique et des différents tempi (225, 600 et 900 ms). Les droites représentent l'estimation (coefficients de régression estimés) de la modélisation et les zones en surbrillances colorées l'erreur standard. longueur(r) : stabilité moyenne de la SMS.

8.2.1.3 SYNTHÈSE DE LA STABILITÉ DE LA SMS

L'exploration descriptive de la stabilité de la SMS a montré une plus grande difficulté à mettre en place la synchronisation pour le tempo de métronome 225 ms par rapport aux tempi 600 et 900 ms (cf. **Tableau 10**), une tendance centrale pour la stabilité de la SMS globale (médiane) de 0,891⁹⁰ ainsi que plusieurs tendances. Le tempo auquel les participants ont tendance à être les plus stables est 900 ms, suivi par le tempo 600 ms puis le tempo 225 ms. La stabilité de la SMS tend à augmenter avec l'âge et le nombre d'années de pratique de la musique et de la danse. La stabilité de la SMS tend à être plus élevée chez les hommes et les musiciens (et notamment les percussionnistes).

Les résultats du premier modèle inférentiel comportant uniquement notre variable explicative d'intérêt « *Tempi_SMS* » montrent une différence de stabilité de la SMS selon le tempo. La stabilité de la SMS est moins élevée pour le tempo 225 ms que pour les tempi 600 et 900 ms, ce qui semble cohérent avec le TMS global de l'échantillon (médiane du groupe = 712 ms, cf. 8.1.1, p. 134). Ce résultat semble aller dans le sens de notre Hypothèse 1.1, selon laquelle le tempo auquel la synchronisation est la plus stable et précise se rapproche de la période de référence (TMS moyen).

En ajustant mieux notre modèle à notre échantillon (ajout de variables explicatives contrôles socio-démographiques), nous trouvons un effet différentiel plus important entre les trois tempi ainsi qu'un effet différentiel entre les tempi lié à l'âge et à l'augmentation du nombre d'années de pratique de la musique. Dans ce second modèle plus ajusté, la stabilité de la SMS est la plus élevée pour le tempo 900 ms et la plus basse pour le tempo 225 ms. L'augmentation du nombre d'années de pratique de la musique est associée à une augmentation de la stabilité pour le tempo 225 ms.

La comparaison des deux modèles (avec et sans variables contrôles socio-démographiques) nous invite à prendre en considération et mieux contrôler pour de futures recherches certaines caractéristiques socio-démographiques telles que l'âge et le nombre d'années de pratique de la musique, qui viennent nuancer les résultats du premier modèle. Cependant, en raison du caractère non contrôlé des variables socio-démographiques dans notre étude, les effets d'interaction entre le tempo et les variables contrôles observés dans le second modèle sont néanmoins à interpréter avec précaution.

⁹⁰ Pour rappel, la stabilité du participant est donnée par un vecteur de longueur moyenne \bar{R} , comprise entre 0 (instable) et 1 (stable).

Nous retrouvons ainsi les valeurs rapportées dans la littérature quant à la stabilité de la pulsation pour la SMS. Canette et al. (2020) rapportent ainsi une stabilité moyenne de la synchronisation sensori-motrice aux alentours de 0,85 (ISI⁹¹ 400 ms = 0,87 ; ISI 550 = 0,84 ; ISI 700 = 0,87). Sowiński et Dalla Bella (2013) rapportent dans leur deuxième étude une valeur de stabilité moyenne pour une SMS à un rythme de métronome de 0,95 (pour les ISI 450, 600 et 750 ms) et pour une SMS à un extrait musical de 0,84 (ISI 450 ms), 0,86 (ISI = 600 ms) et 0,88 (ISI = 750 ms).

Nous retrouvons également les observations de Drake et al. (2000) ou Repp (2005), qui rapportent une incidence du niveau d'expertise musical sur la SMS, les participants musiciens ayant plus de facilité à se synchroniser à un rythme extérieur que les non-musiciens.

Ces résultats concernant la stabilité de la SMS seront ensuite mis en regard des observations concernant la tâche de Tempo Moteur Spontané (TMS) (Activité 1) et l'Activité 2 afin d'investiguer l'*Hypothèse 6*, à propos de l'existence de mécanismes cognitifs partagés pour le traitement du rythme et du langage. En effet, nous nous attendons à ce que les participants les plus stables et précis dans les tâches de TMS et SMS soient ceux qui présentent le moins d'erreurs de lecture et un Temps d'Initiation (TI) de la parole plus rapide. Nous reviendrons sur ces comparaisons entre tâches dans la partie 10 (p. 187).

⁹¹ De l'anglais *Inter-Stimuli Interval*.

8.2.2 PRECISION DE LA SYNCHRONISATION SENSORI-MOTRICE

Tout comme la stabilité du Tempo Moteur Spontané (TMS) et de la SMS, la précision de la SMS est également calculée à partir des statistiques circulaires (cf. 6.2.2, p. 115). Elle est indiquée par l'angle ($\bar{\theta}$) du vecteur circulaire, représentatif de l'asynchronie moyenne entre le battement de métronome entendu et la frappe du participant. Nous nous intéressons plus précisément à la valeur absolue de $\bar{\theta}$ et à son signe (positif ou négatif). La valeur absolue de $\bar{\theta}$ est comprise entre 0° (en parfaite synchronie avec le battement de métronome) et 180° (à contre-temps du battement de métronome). Une valeur absolue proche de 0° témoigne ainsi d'une haute précision de la SMS. Le signe de l'angle $\bar{\theta}$ indique si les frappes du participant sont en avance (pour un signe négatif) ou en retard (pour un signe positif) par rapport au battement de métronome entendu. Une anticipation témoigne de la mise en place de la synchronisation, tandis qu'un retard reflète plutôt un comportement de réaction au stimulus (Lenc et al., 2021).

Les participants étaient exposés à chaque tempo sur deux blocs expérimentaux (cf. 3.1.2, p. 87), nous avons ainsi recueilli deux valeurs moyenne par participant et par tempo pour la mesure de la précision de la SMS.

Nous présentons les résultats (statistiques descriptives et inférentielles) de la valeur absolue de la précision de la SMS en partie 8.2.2.1 et du signe de la précision de la SMS en partie 8.2.2.2 (p. 161).

8.2.2.1 VALEUR ABSOLUE DE L'ANGLE $\bar{\theta}$

8.2.2.1.1 EXPLORATION DESCRIPTIVE

Les valeurs absolues d'angles $\bar{\theta}$, représentatives de la précision ou de l'asynchronie de la SMS, ont été calculées à partir de la moyenne des deux blocs expérimentaux de chaque participant, pour chaque tempo, à condition que la stabilité moyenne du participant satisfasse le test de Rayleigh⁹². Nous avons ensuite agrégé les moyennes d'angle des trois tempi pour calculer la précision moyenne des participants tous tempi confondus ($n = 38^{93}$).

⁹² Pour rappel, le test de Rayleigh permet de distinguer la répartition des frappes du participant d'une répartition aléatoire, et témoigne de la réussite de la mise en place de la synchronisation.

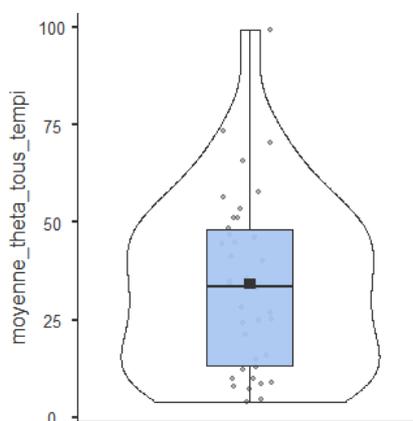
⁹³ 40 participants * 1 observation, sans les valeurs manquantes.

A partir de l'ensemble des valeurs de précisions moyennes agrégées de chaque participant tous tempi confondus ($n = 38$), nous avons pu calculer la précision de la SMS moyenne globale, moyenne des valeurs d'angles $\bar{\theta}$ individuelles de notre échantillon. La distribution des valeurs absolues d'angles $\bar{\theta}$ des individus suivant une distribution normale ($p = 0,05$ au test de normalité de Shapiro-Wilk, cf. **Figure 46** et **Tableau 12**), nous avons calculé la moyenne comme indice de tendance centrale.

La précision globale de la SMS pour l'ensemble de notre échantillon tous tempi confondus est de $\bar{\theta} = 34,1^\circ$ (valeur absolue).

Figure 46 :

Représentation graphique de la précision de la SMS des participants



Note. La boîte à moustaches indique la médiane et les quartiles 1 et 3 ; le carré la moyenne ; le violon la densité. moyenne_theta_tous_tempi : précision moyenne de la SMS, tous tempi confondus.

Tableau 12 :

Valeurs descriptives de la précision de la SMS des participants

| Statistiques descriptives | | | | | | | | Shapiro-Wilk | |
|---------------------------|----|-----------|---------|---------|------------|---------|---------|--------------|-------|
| | N | Manquants | Moyenne | Médiane | Ecart-type | Minimum | Maximum | W | p |
| moyenne_theta_tous_tempi | 38 | 2 | 34.1 | 33.5 | 22.5 | 3.87 | 99.4 | 0.942 | 0.050 |

Note. Valeurs descriptives de la précision de la SMS des participants (valeurs absolues d'angle theta), tous tempi confondus. Un $p < 0,05$ au test de Shapiro-Wilk indique une distribution non paramétrique des données. moyenne_theta_tous_tempi : précision moyenne de la SMS, tous tempi confondus.

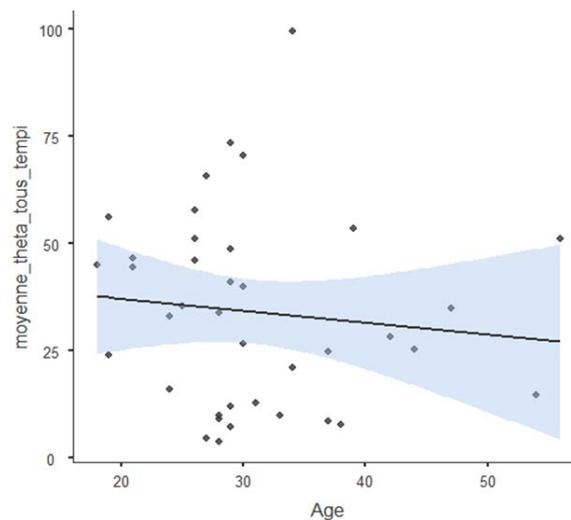
De plus, en explorant les données, nous observons plusieurs tendances concernant la précision de la SMS, selon le tempo du métronome (ISI 225, 600 ou 900 ms). Le tempo auquel les participants sont les plus précis (valeur de $\bar{\theta}$ la plus basse) est 900 ms (médiane = 24,9°), suivi par le tempo 600 ms (médiane = 25,7°) et le tempo 225 ms (médiane = 26,5°)

Pour rappel, nous avons également souhaité observer l'effet des variables socio-démographiques sur la précision de la SMS, et avons ainsi intégré dans nos analyses les observations concernant l'effet de l'âge, du sexe, de la pratique musicale, de la danse et des jeux vidéo (Oui/Non) et du nombre d'années de pratique de la danse et de la musique des participants⁹⁴. Notre exploration descriptive semble ainsi indiquer plusieurs tendances.

La précision augmente (la valeur de l'angle diminue) avec l'avancée de l'âge, du nombre d'années de pratique de la musique et de la danse (cf. **Figure 47**, **Figure 48** et **Figure 49**)

Figure 47 :

Evolution de la précision de la SMS en fonction de l'âge

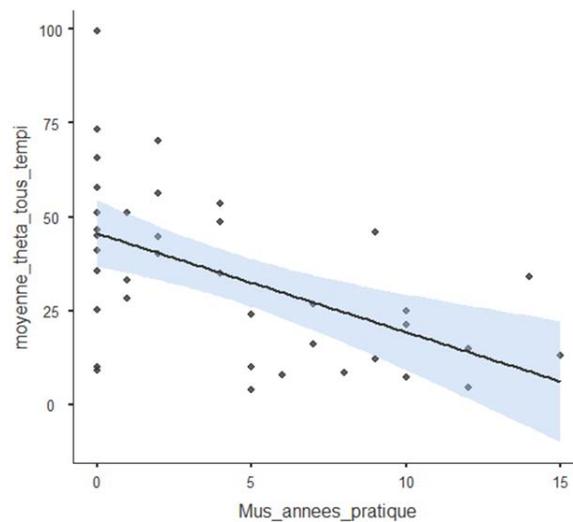


Note. Evolution de la précision de la SMS (valeurs absolues d'angles, en degrés) en fonction de l'âge. La ligne indique la droite de régression linéaire et l'espace coloré l'erreur standard. moyenne_theta_tous_tempi : précision moyenne de la SMS, tous tempi confondus.

⁹⁴ Pour rappel, dans cette étude, nous employons les termes « musicien », « danseur » ou « joueur de jeux vidéo » dès lors qu'un participant a indiqué pratiquer ou avoir pratiqué une activité musicale, de danse ou de jeux vidéo respectivement.

Figure 48 :

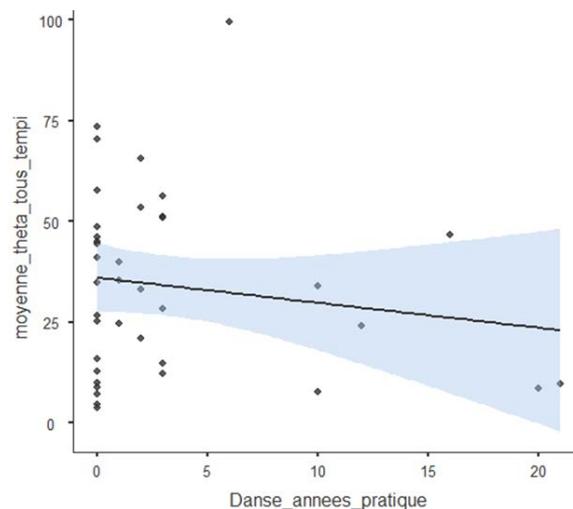
Evolution de la précision de la SMS en fonction du nombre d'années de musique



Note. Evolution de la précision de la SMS (valeurs absolues d'angles, en degrés) en fonction du nombre d'années de pratique de la musique. La ligne indique la droite de régression linéaire et l'espace coloré l'erreur standard. moyenne_theta_tous_tempi : précision moyenne de la SMS, tous tempi confondus.

Figure 49

Evolution de la précision de la SMS en fonction du nombre d'années de danse



Note. Evolution de la précision de la SMS (valeurs absolues d'angles, en degrés) en fonction du nombre d'années de pratique de la danse. La ligne indique la droite de régression linéaire et l'espace coloré l'erreur standard. moyenne_theta_tous_tempi : précision moyenne de la SMS, tous tempi confondus.

De plus, il semblerait que la précision soit plus élevée (la valeur de l'angle est la plus basse) chez les hommes (moyenne = 27°) par rapport aux femmes (moyenne = $40,4^\circ$) ; que la précision soit plus élevée (la valeur de l'angle est la plus basse) chez les participants musiciens (moyenne = $28,3^\circ$) par rapport aux non-musiciens (moyenne = $46,61^\circ$) ; et que la précision soit la plus élevée chez les musiciens percussionnistes (moyenne = $5,55^\circ$) et la moins élevée chez les non-musiciens (moyenne = $46,61^\circ$).

8.2.2.1.2 ANALYSE INFÉRENTIELLE

Afin d'investiguer l'*Hypothèse 1* et de modéliser la relation entre le tempo et la précision de la SMS, nous avons ensuite appliqué un modèle linéaire mixte à nos données de précision ($n = 227^{95}$).

Dans un premier temps, nous avons appliqué à nos données un modèle comprenant uniquement notre variable indépendante d'intérêt (Tempi_SMS) (cf. 8.2.2.1.2.1). Dans un second temps, nous avons construit et appliqué un second modèle incluant également des variables contrôles socio-démographiques afin d'estimer la potentielle interaction entre les variables socio-démographiques et les effets d'intérêt (cf. 8.2.2.1.2.2, p. 160).

8.2.2.1.2.1 EFFET DU TEMPO SUR LA PRÉCISION DE LA SMS

Afin de modéliser la relation entre le tempo et la précision moyenne de la SMS des participants, nous avons appliqué un modèle linéaire mixte à nos données de précision ($n = 227$), avec en effets fixes le tempo, en effet aléatoire les participants et en variable dépendante la précision de la SMS.

Après avoir établi la structure aléatoire la plus ajustée à notre modèle (cf. ANNEXES – PARTIE 2, document complémentaire, 1.1.2.2, p. 51), nous avons appliqué celui-ci à nos données de précision de la SMS. Nous n'avons pas opéré de sélection de modèle pour les effets fixes, puisque nous n'avons qu'un seul paramètre fixe : le tempo.

Il ressort de la modélisation que les différences de tempo (ISI 225, 600 et 900 ms) sur la précision de la SMS ne sont pas significatives ($p = 0,065$).

8.2.2.1.2.2 EFFET DU TEMPO ET DES CARACTÉRISTIQUES SOCIO-DEMOGRAPHIQUES SUR LA PRÉCISION DE LA SMS

Après avoir estimé l'effet du tempo (variable d'intérêt) sur la précision de la SMS, nous avons voulu aller plus loin en ajoutant au modèle des caractéristiques socio-démographiques (variables contrôles) qui permettraient de contrôler si la variance observée est liée à une interaction entre les variations de tempo et les caractéristiques socio-démographiques des participants.

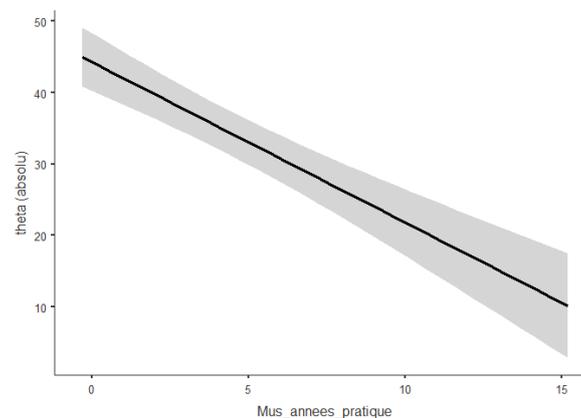
⁹⁵ Données réelles/désagrégées : 40 participants * 3 tempi * 2 blocs, sans les valeurs manquantes.

Nous sommes partie d'un modèle saturé avec les paramètres suivants : facteur d'intérêt tempo ; covariables contrôles âge, années de pratique de la musique et de la danse ; facteurs contrôles sexe, pratique de la musique, de la danse et des jeux vidéo ; et les interactions doubles entre d'une part le facteur d'intérêt et d'autre part les variables contrôles. Après avoir comparé pas à pas la pertinence des paramètres pour parvenir à un modèle le plus simple et le plus ajusté (cf. ANNEXES – PARTIE 2, document complémentaire, 1.1.2.2, p. 54), nous avons appliqué le nouveau modèle à nos données.

Nous n'observons pas d'effet significatif du tempo (ISI 225, 600 et 900 ms) sur la précision de la SMS, simple ($p = 0,065$) ou en interaction avec les caractéristiques socio-démographiques. En revanche, nous observons un effet du nombre d'années de pratique de la musique sur la variabilité de la précision de la SMS. La précision de la SMS augmente (la valeur de l'angle diminue) avec l'augmentation du nombre d'années de pratique de la musique ($Estimate = -2,25$; $p < 0,001$) (cf. **Figure 50**).

Figure 50 :

Evolution de la précision de la SMS en fonction du nombre d'années de musique



Note. Evolution de la précision de la SMS (valeurs absolues d'angles, en degrés) en fonction du nombre d'années de pratique de la musique. La droite représente l'estimation (coefficient de régression estimé) de la modélisation, les zones en surbrillances colorée l'erreur standard. theta (absolu) : précision moyenne de la SMS.

8.2.2.2 SIGNE DE THETA

Le signe de $\bar{\theta}$ est majoritairement négatif pour les tempi 600 ms (dans 94,94 % des cas, $n = 79^{96}$) et 900 ms (dans 92,5 % des cas, $n = 80^{97}$), ce qui indique que les participants anticipent le battement de métronome, et sont donc synchronisés à sa fréquence. En revanche, pour le tempo 225 ms, la répartition du signe de $\bar{\theta}$ est plus équilibrée (négatif dans 57,35 % des cas et

⁹⁶ 40 participants * 2 blocs expérimentaux, sans une valeur aberrante (longueur \bar{R} non significative au test de Rayleigh).

⁹⁷ 40 participants * 2 blocs expérimentaux.

positif 42,65 % des cas, $n = 68^{98}$). Le signe positif de $\bar{\theta}$ témoigne d'une réaction au stimulus (à la différence d'un signe négatif indiquant une anticipation) et donc d'une plus grande difficulté à établir la synchronisation au tempo de 225 ms par rapport aux tempi 600 ms et 900 ms (cf. **Tableau 13**). Cette observation est cohérente avec les limites de la synchronisation sensori-motrice (cf. 1.2.1.1), qui devient plus difficile lorsque le tempo d'un stimulus auditif dépasse le seuil de 200-250 ms (Repp, 2003).

Tableau 13 :

Tableau de fréquences du signe de $\bar{\theta}$ (moins/plus) en fonction du tempo

| Fréquences de theta (signe) | | | | |
|-----------------------------|-----------|-----------|------------|-----------|
| theta (signe) | Tempi_SMS | Quantités | % du Total | % cumulés |
| moins | 225_ms | 39 | 17.2 % | 17.2 % |
| | 600_ms | 75 | 33.0 % | 50.2 % |
| | 900_ms | 74 | 32.6 % | 82.8 % |
| plus | 225_ms | 29 | 12.8 % | 95.6 % |
| | 600_ms | 4 | 1.8 % | 97.4 % |
| | 900_ms | 6 | 2.6 % | 100.0 % |

Note. Le signe de $\bar{\theta}$ indique l'asynchronie positive ou négative moyenne des participants. Une asynchronie positive témoigne d'une réaction aux battements de métronome, à la différence d'une asynchronie négative indiquant une anticipation et donc la mise en place de la synchronisation.

8.2.2.3 SYNTHÈSE DE LA PRÉCISION DE LA SMS

Concernant la valeur absolue de l'angle du vecteur circulaire, l'exploration descriptive de la précision de la SMS a montré une tendance centrale pour la précision globale (moyenne) de 34,1° ainsi que plusieurs tendances. Le tempo auquel les participants ont tendance à être les plus précis est 900 ms, suivi par le tempo 600 ms puis le tempo 225 ms. La précision de la SMS tend à augmenter avec l'âge et le nombre d'années de pratique de la musique et de la danse. La précision de la SMS tend à être plus élevée chez les hommes et les musiciens (et notamment les percussionnistes). Ces tendances semblent aller dans le sens de l'*Hypothèse 1.1*, indiquant une plus grande précision de la SMS pour un tempo se rapprochant du TMS moyen. Le TMS moyen global de notre échantillon est 712 ms, les tempi de SMS s'en rapprochant le plus sont donc 600 et 900 ms, tandis que la précision au tempo 225 ms est plus faible.

Les résultats du premier modèle inférentiel comportant uniquement notre variable explicative d'intérêt (tempo) ne montrent pas de différence significative de précision de la SMS

⁹⁸ 40 participants * 2 blocs expérimentaux, sans les valeurs aberrantes (longueurs \bar{R} non significatives au test de Rayleigh).

selon le tempo, et ne permettent ainsi pas d'appuyer l'*Hypothèse 1.1*, selon laquelle le tempo auquel la synchronisation est la plus stable et précise se rapproche de la période de référence (TMS moyen).

Le second modèle, plus ajusté à notre échantillon par l'ajout de variables contrôles socio-démographiques, n'indique pas non plus de différence dans la précision de la SMS selon le tempo, ni en effet simple, ni en interaction avec des variables contrôles socio-démographiques. En revanche, la précision de la SMS augmente (la valeur d'angle diminue) significativement avec l'augmentation du nombre d'années de pratique de la musique.

La comparaison des deux modèles inférentiels nous invite à prendre en considération et mieux contrôler pour de futures recherches le nombre d'années de pratique de la musique, qui vient nuancer les résultats du premier modèle. Cependant, en raison du caractère non contrôlé des variables socio-démographiques dans notre étude, les effets d'interaction entre le tempo et les variables contrôles sont néanmoins à interpréter avec précaution.

Concernant le signe de theta, l'exploration descriptive montre que la SMS tend à être mieux établie pour les tempi 600 et 900 ms par rapport au tempo 225 ms, ce qui semble aller dans le sens de notre *Hypothèse 1.1*, selon laquelle le tempo auquel la synchronisation s'établit plus facilement à un tempo de SMS qui se rapproche de la période de référence (TMS moyen), soit les tempi de métronome 600 et 900 ms.

Nous retrouvons les valeurs rapportées dans la littérature quant à la précision de la pulsation pour la SMS. Canette et al. (2020) rapportent ainsi une précision moyenne de la synchronisation sensori-motrice aux alentours de 50° (ISI⁹⁹ 400 ms = $-46,09^\circ$; ISI 550 = $-72,80^\circ$; ISI 700 = $-51,70^\circ$). Sowiński et Dalla Bella (2013) rapportent dans leur deuxième étude une valeur de précision moyenne pour une SMS à un rythme de métronome de $-26,5^\circ$ (ISI = 450 ms), de $-22,9^\circ$ (ISI = 600 ms) et de $-20,9^\circ$ (ISI = 750 ms), ainsi qu'une valeur de précision moyenne pour une SMS à un extrait musical de $4,1^\circ$ (ISI 450 ms), $-5,2^\circ$ (ISI = 600 ms) et $-8,9^\circ$ (ISI = 750 ms).

⁹⁹ De l'anglais *Inter-Stimuli Interval*.

Nous retrouvons également les observations de Drake et al. (2000) ou Repp (2005), qui rapportent une incidence du niveau d'expertise musical sur la SMS, les participants musiciens ayant plus de facilité à se synchroniser à un rythme extérieur que les non-musiciens.

Ces résultats concernant la précision de la SMS seront ensuite mis en regard des observations concernant la tâche de Tempo Moteur Spontané (TMS) (Activité 1) et l'Activité 2 afin d'investiguer l'*Hypothèse 6*, à propos de l'existence de mécanismes cognitifs partagés pour le traitement du rythme et du langage. En effet, nous nous attendons à ce que les participants les plus stables et précis dans les tâches de TMS et SMS soient ceux qui présentent le moins d'erreurs de lecture et un Temps d'Initiation (TI) de la parole plus rapide. Nous reviendrons sur ces comparaisons entre tâches dans la partie 10 (p. 187).

8.2.3 TEMPO LE PLUS STABLE ET/OU LE PLUS PRÉCIS

En comparant les vecteurs moyens de chaque tempo pour chaque participant ($n = 40 * 3$ tempi), nous avons pu faire ressortir le tempo pour lequel la synchronisation a été la mieux réussie pour chaque participant, pour la stabilité et la précision. Concernant la stabilité, le tempo le mieux réussi correspond au tempo pour lequel la valeur de la longueur \bar{R} est la plus élevée, tandis que pour la précision, le tempo le mieux réussi correspond au tempo pour lequel la valeur absolue de l'angle $\bar{\theta}$ est la plus basse¹⁰⁰. Par exemple, dans le Tableau 14, le tempo auquel la SMS est la plus stable pour ce participant est 900 ms, associé à la valeur de stabilité la plus élevée (0,963), et le tempo auquel la SMS est la plus précise est 600 ms, associé à la valeur de précision la plus basse (37,9°).

Tableau 14 :

Description de la stabilité et de la précision de la SMS d'un participant

Statistiques descriptives

| | Stabilité 225 ms | Stabilité 600 ms | Stabilité 900 ms | Tempo le + stable | Précision 225 ms | Précision 600 ms | Précision 900 ms | Tempo le + précis |
|---------|---------------------|---------------------|---------------------|-------------------------|---------------------|---------------------|---------------------|-------------------------|
| Moyenne | 0.499 | 0.936 | 0.963 | 900 ms | 59.5 | 37.9 | 42.6 | 600 ms |

Note. Exemple des valeurs de stabilité et précision selon le tempo, permettant d'identifier le tempo le plus stable et le tempo le plus précis, ici pour un participant.

Le tempo de 225 ms comme tempo le plus stable n'a été retrouvé chez aucun participant, tandis que le tempo de 900 ms comme tempo le plus stable est le plus répandu (cf. **Tableau 15**).

Tableau 15 :

Tableau de fréquences des tempi les plus stables lors de la SMS

Fréquences de longueur_max

| longueur_max | Quantités | % du Total | % cumulés |
|--------------|-----------|------------|-----------|
| 225_ms | 0 | 0 % | 0 % |
| 600_ms | 13 | 32.5 % | 32.5 % |
| 900_ms | 27 | 67.5 % | 100.0 % |

Note. N = 40 participants.

¹⁰⁰ Pour rappel, la valeur absolue de l'angle $\bar{\theta}$ renvoie à l'asynchronie. Ainsi, lorsque la valeur de l'angle est basse, l'asynchronie est basse, donc la précision est élevée.

Concernant la précision de la SMS, les trois tempi sont équitablement répartis en tant que tempo le plus précis (cf. **Tableau 16**).

Tableau 16 :

Tableau de fréquences des tempi les plus précis de la SMS

Fréquences de theta_min

| theta_min | Quantités | % du Total | % cumulés |
|-----------|-----------|------------|-----------|
| 225_ms | 13 | 32.5 % | 67.5 % |
| 600_ms | 13 | 32.5 % | 100.0 % |
| 900_ms | 14 | 35.0 % | 35.0 % |

Note. N = 40 participants.

Nous observons de plus une incidence du profil musical sur le tempo le plus stable et/ou le plus précis. La répartition des tempi les plus stables est plus équilibrée chez les participants pratiquant ou ayant pratiqué une activité musicale (cf. **Tableau 17**), par rapport aux non-musiciens (cf. **Tableau 18**).

Tableau 17 :

Tableau de fréquences des tempi les plus stables lors de la SMS chez les musiciens

Fréquences de longueur_max

| longueur_max | Quantités | % du Total | % cumulés |
|--------------|-----------|------------|-----------|
| 600_ms | 12 | 44.4 % | 44.4 % |
| 900_ms | 15 | 55.6 % | 100.0 % |

Note. N = 27 participants musiciens.

Tableau 18 :

Tableau de fréquences des tempi les plus stables lors de la SMS chez les non-musiciens

Fréquences de longueur_max

| longueur_max | Quantités | % du Total | % cumulés |
|--------------|-----------|------------|-----------|
| 600_ms | 1 | 7.7 % | 7.7 % |
| 900_ms | 12 | 92.3 % | 100.0 % |

Note. N = 13 participants non-musiciens

La répartition des tempi les plus précis est également plus équilibrée chez les participants musiciens (cf. **Tableau 20**) par rapport aux non-musiciens (cf. Tableau 20).

Tableau 19 :

Tableau de fréquences des tempi les plus précis lors de la SMS chez les musiciens

Fréquences de theta_min

| theta_min | Quantités | % du Total | % cumulés |
|-----------|-----------|------------|-----------|
| 225_ms | 11 | 40.7 % | 40.7 % |
| 600_ms | 7 | 25.9 % | 66.7 % |
| 900_ms | 9 | 33.3 % | 100.0 % |

Note. N = 27 participants musiciens.

Tableau 20 :

Tableau de fréquences des tempi les plus précis lors de la SMS chez les non-musiciens

Fréquences de theta_min

| theta_min | Quantités | % du Total | % cumulés |
|-----------|-----------|------------|-----------|
| 225_ms | 2 | 15.4 % | 15.4 % |
| 600_ms | 6 | 46.2 % | 61.5 % |
| 900_ms | 5 | 38.5 % | 100.0 % |

Note. N = 13 participants non-musiciens

Ces observations rejoignent ainsi celles de Drake et al. (2000) à propos de l'effet de l'expertise musicale sur l'étendue des tempi de synchronisation sensori-motrice, les participants musiciens ayant une plus grande facilité que les non-musiciens à se synchroniser sur une plus grande étendue de tempi. A nouveau, nous rappelons cependant que les informations socio-démographiques n'ont pas été spécifiquement contrôlées dans cette étude. Les participants musiciens sont ceux ayant indiqué pratiquer ou avoir pratiqué une activité musicale.

Les résultats concernant la répartition des tempi les plus stables et précis sur l'ensemble des participants sont ainsi cohérents avec les résultats des analyses inférentielles concernant l'effet du tempo sur la stabilité et la précision de la SMS. De plus, les observations à propos de la répartition du tempo le plus stable sont cohérentes avec l'*Hypothèse 1.1*, selon laquelle le tempo le plus stable se rapproche de la période de référence (TMS moyen). En effet, nous ne

retrouvons pour aucun participant un tempo le plus stable à 225 ms, qui est le tempo le plus éloigné de la période de référence (TMS moyen global de l'échantillon = 712 ms). En revanche, les observations à propos de la répartition du tempo le plus précis ne vont pas dans ce sens, les trois tempi étant aussi bien réussis du point de vue de la précision. Nous reviendrons sur la comparaison des tâches de TMS, de SMS et de l'Activité 2 dans la partie 10 (p. 187).

9 ACTIVITE 2 : AMORÇAGE RYTHMIQUE DE LA PAROLE

Lors de l'Activité 2, les participants ont pour consigne de lire des syntagmes à haute voix, chacun des syntagmes étant amorcé par un stimulus rythmique. Nous avons recueilli les erreurs de lecture produites ainsi que le Temps d'Initiation (TI) de la parole, ces deux variables dépendantes faisant office d'indice quant à la facilité des participants à produire de la parole (Shao & Meyer, 2018).

L'objectif principal de cette activité est d'observer et comparer l'effet de différents amorçages rythmiques sur la facilitation de la production de la parole. Plus particulièrement, nous nous sommes intéressée à l'effet de trois variables indépendantes sur la production de la parole : l'Engagement Moteur (EM) du participant (avec ou sans *tapping*), le Type d'Amorce (TA) (langagier, non langagier ou musical) et la Congruence Métrique (CM) (structure métriquement congruente ou incongruente), renvoyant à nos Hypothèse 2 à 5 (cf. p. 79). Les trois VI (EM, TA et CM) étant croisées, nous avons ainsi un plan expérimental en 12 conditions : une amorce est ainsi caractérisée à la fois par l'engagement moteur du participant (avec +EM ou sans -EM), le type d'amorce (vocale langagière L, vocale non langagière NL ou musicale M) et la congruence métrique de l'amorçage (congruent C ou incongruent IC) (cf. 3.2, p. 89). Nous avons recueilli 8 syntagmes cibles pour chacune des 12 conditions expérimentales, soit un total de 96 syntagmes recueillis par participant et un total de 3840 syntagmes pour l'ensemble des 40 participants.

Pour rappel, cette activité se situe dans la continuation des recherches sur l'effet de l'amorçage rythmique sur le traitement du langage. Nous poursuivons plus particulièrement les travaux engagés par Cason, Astésano, et al. (2015) dont nous réutilisons une partie du matériel expérimental. Dans cette optique, notre travail se donne pour objectif une réplique partielle des observations de Cason, Astésano, et al. (2015). En effet, les auteurs ont trouvé qu'un amorçage métriquement congruent sur le versant de la perception facilite la détection de phonèmes par rapport à un amorçage incongruent (incongruence partielle ou complète¹⁰¹), et ce plus particulièrement pour un groupe de participants ayant suivi un court exercice d'engagement moteur (copie vocale de l'amorce). Les auteurs rapportent également des effets différentiels de la congruence métrique de l'amorçage, selon la structure de l'amorce. Nous

¹⁰¹ Incongruence métrique partielle : la structure métrique de l'amorce diffère de celle de la cible par un seul paramètre (emplacement du 1^{er} temps métriquement fort ou nombre d'éléments) ; incongruence métrique complète : la structure métrique de l'amorce diffère de celle de la cible à la fois en ce qui concerne l'emplacement du 1^{er} temps fort et en nombre d'éléments). Pour un rappel des 4 conditions de congruence de l'étude Cason, Astésano, et al. (2015), se rapporter à la **Figure 13** (p. 74).

adaptions ce protocole cette fois pour le versant de la production, en y ajoutant l'objectif principal de faire émerger un effet facilitateur du Type d'Amorce (TA) sur la production de la parole. De plus, notre condition d'Engagement Moteur (EM) se distingue également de celle de Cason, Astésano, et al. (2015) en ce qu'elle utilise la motricité manuelle et non articulaire. Afin de ne pas surcharger notre plan expérimental, nous avons choisi de ne pas conserver la granularité en quatre niveaux de congruence métrique utilisée dans Cason et al. (2015) (1. congruent ; 2. et 3. partiellement incongruent ; 4. complètement incongruent, cf. **Figure 13**, p. 76) et avons sélectionné uniquement les conditions d'amorçage métriquement congruent (1.) et complètement incongruent (4.).

La mesure du Temps d'Initiation (TI) de la parole, mise en regard des mesures de TMS moyen et de tempo le plus stable et/ou le plus précis (Activité 1) dans la partie 10, p. 187, nous permet également d'investiguer l'Hypothèse 1, à propos d'une période de référence supramodale. De plus, les mesures de nombre d'erreurs de lecture et de TI de la parole, représentatives des capacités de traitement du langage, sont mises en regard des capacités de traitement du rythme (stabilité et précision du TMS et de la SMS) dans la partie 10, p. 187, afin d'investiguer l'Hypothèse 6.

De même que nous avons observé l'effet des variables socio-démographiques sur les mesures de l'Activité 1, nous avons pensé intéressant d'explorer les liens possibles entre nos deux mesures (nombre d'erreurs de lecture et Temps d'Initiation de la parole) et les caractéristiques socio-démographiques des participants (données du questionnaire, cf. 6.1, p. 113). Nous avons donc intégré les observations concernant l'âge, le sexe, la pratique musicale, de la danse et des jeux vidéo (Oui/Non) et le nombre d'années de pratique de la danse et de la musique dans les analyses descriptives et inférentielles.

9.1 ERREURS DE LECTURE

Nous avons relevé les erreurs de lecture relevant des catégories suivantes : répétitions ou reformulations ; substitutions grammaticales, lexicales, phonétiques, phonologiques, de genre et dérivationnelles ; les syntagmes avec omissions phonologiques, morphologiques ou lues partiellement ; les syntagmes avec déplacement phonologique ; les syntagmes avec ajouts de phones et lexicaux (cf. 6.3.2.1, p. 120). Les erreurs peuvent ainsi être le reflet d'une interférence avec le bon déroulé des étapes d'encodage lexical et d'articulation, concernant les étapes séquentielles ou bien les boucles de monitoring ou de contrôle anticipatoire (cf. les modèles neuropsycholinguistiques de production de la parole en partie 2.1.2, p. 42).

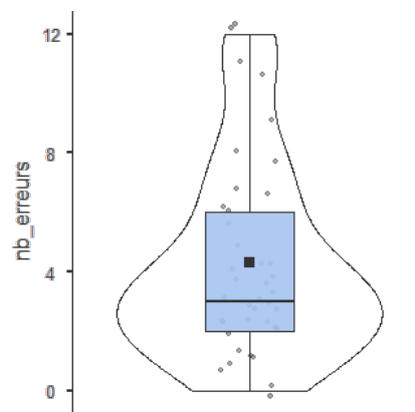
9.1.1 EXPLORATION DESCRIPTIVE

A partir de l'ensemble des nombres d'erreurs de chaque participant toutes conditions confondues ($n = 40$), nous avons pu calculer le nombre d'erreurs global, et la valeur médiane des nombres d'erreurs individuels de notre échantillon. La distribution des nombres d'erreurs des individus ne suivant pas une distribution normale ($p < 0,001$ au test de normalité de Shapiro-Wilk, cf. **Figure 51** et **Tableau 21**), nous avons calculé la médiane comme indice de tendance centrale.

Le nombre moyen d'erreurs de lecture global sur l'ensemble de notre échantillon est de 3 erreurs par participant. Sur l'ensemble des syntagmes cibles lus recueillis ($n = 3840^{102}$), 172 ont présenté des erreurs de lecture, soit 4,48 % du total des syntagmes lus.

Figure 51 :

Représentation graphique du nombre d'erreurs de lecture moyen des participants



Note. La boîte à moustaches indique la médiane et les quartiles 1 et 3 ; le carré la moyenne ; le violon la densité.

¹⁰² 40 participants * 96 syntagmes cibles.

Tableau 21 :*Valeurs descriptives du nombre d'erreurs de lecture moyen des participants*

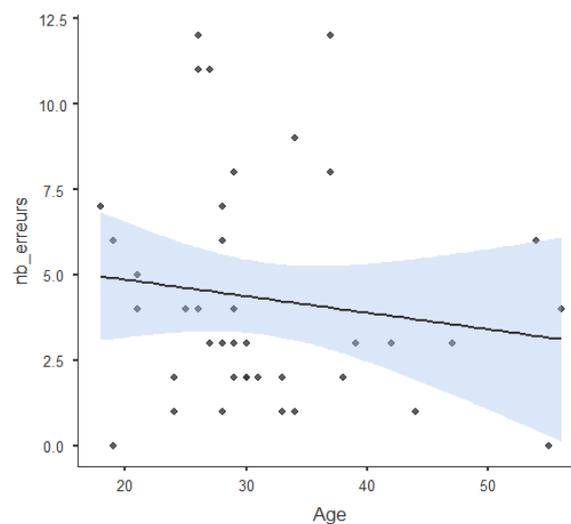
Statistiques descriptives

| | N | Moyenne | Médiane | Ecart-type | Minimum | Maximum | Shapiro-Wilk | |
|------------|----|---------|---------|------------|---------|---------|--------------|--------|
| | | | | | | | W | p |
| nb_erreurs | 40 | 4.30 | 3.00 | 3.29 | 0 | 12 | 0.887 | < .001 |

Note. Un $p < 0,05$ au test de Shapiro-Wilk indique une distribution non paramétrique des données.

Concernant nos trois variables indépendantes (Engagement Moteur, Type d'Amorce et Congruence Métrique), les participants ont fait plus d'erreurs de lecture en condition +EM (63,4 %) qu'en condition -EM (36,6 %), plus d'erreurs de lecture en condition M (amorce Musicale) (37,8 %) qu'en condition NL (amorce vocale Non Langagière) (31,4 %) et L (amorce vocale Langagière) (30,8 %) et enfin plus d'erreurs de lecture en situation d'amorçage congruent (54,7 %) qu'incongruent (45,3 %).

De plus, nous observons une tendance à la diminution du nombre d'erreurs avec l'avancée de l'âge (cf. **Figure 52**).

Figure 52 :*Evolution du nombre d'erreurs de lecture en fonction de l'âge*

Note. La ligne indique la droite de régression linéaire et l'espace coloré l'erreur standard.

9.1.2 ANALYSE INFÉRENTIELLE

Afin de répondre à nos *Hypothèse 2 à 5* et de modéliser la relation entre les différences d'amorçage rythmique (Engagement Moteur – EM, Type d'Amorce – TA et Congruence Métrique – CM) et la production d'erreurs de lecture, nous avons ensuite appliqué un modèle linéaire généralisé à effets mixtes binomial à nos données ($n = 3840^{103}$), avec en effet fixe les différences d'amorçage rythmique (EM, TA et CM), en effet aléatoire les participants et les cibles et en variable dépendante les erreurs produites (coté « 0 » pour un syntagme produit sans erreur de lecture ou « 1 » pour un syntagme comportant au moins une erreur).

Après avoir établi la structure aléatoire la plus ajustée à notre modèle (cf. ANNEXES – PARTIE 2, document complémentaire, 1.2.1, p. 68), nous avons appliqué celui-ci à nos données de production d'erreurs de lecture.

Nous observons un effet principal de l'Engagement moteur ($Exp(B)^{104} = 1,865$; $p < 0,001$) sur la production d'erreurs de lecture. La probabilité de faire des erreurs en condition +EM est donc 1,865 fois plus élevée qu'en condition -EM.

De plus, l'effet principal de EM est modulé par le Type d'Amorce (TA) ($\chi^2(2)=6,478$; $p = 0,039$). L'effet de l'engagement moteur (EM) n'est significatif que pour les TA langagiers (L) ($Exp(B)=2,32$ [1,263 ; 4,25] ; $SE = 0,309$; $p = 0,007$) et non langagiers (NL) ($Exp(B)=2,70$ [1,449 ; 5,03] ; $SE = 0,317$; $p = 0,002$), tandis que cet effet n'est pas significatif pour le TA musical (M) ($p = 0,890$).

En revanche, nous n'observons pas d'effet simple des paramètres « Type_ amorce » ($p = 0,352$) et « Congruence » ($p = 0,428$) ou des interactions « Engagement_moteur : Congruence » ($p = 0,328$) et « Type_ amorce : Congruence » ($p = 0,131$).

A partir des probabilités estimées (moyennes marginales) par le modèle, l'amorçage rythmique pour lequel la probabilité de réaliser au moins une erreur de lecture est la plus importante est l'amorçage avec engagement moteur et type d'amorce non langagier (probabilité +EM_NL = 0,042). Par contraste, l'amorçage rythmique pour lequel la probabilité de réaliser au moins une erreur de lecture est la plus basse est l'amorçage sans engagement moteur non langagier (probabilité -EM_NL = 0,016) (cf. **Tableau 22**).

¹⁰³ 40 participants * 12 conditions expérimentales * 8 observations.

¹⁰⁴ « $Exp(B)$ » est équivalent à l'*odd ratio* (OR).

Tableau 22 :
Probabilités estimées d'apparition d'une erreur selon EM et TA

Estimate Marginal Means - Engagement_moteur * Type_amorce

| Engagement_moteur | Type_amorce | Mean | SE | 95% Confidence Intervals | |
|-------------------|-------------|--------|---------|--------------------------|--------|
| | | | | Lower | Upper |
| +EM | NL | 0.0422 | 0.00910 | 0.02753 | 0.0641 |
| +EM | L | 0.0406 | 0.00880 | 0.02648 | 0.0618 |
| +EM | M | 0.0345 | 0.00796 | 0.02190 | 0.0540 |
| -EM | M | 0.0333 | 0.00776 | 0.02103 | 0.0524 |
| -EM | L | 0.0180 | 0.00517 | 0.01019 | 0.0314 |
| -EM | NL | 0.0161 | 0.00480 | 0.00891 | 0.0287 |

Note. Expected means are expressed as probabilities

Note. Probabilités (colonne Mean) estimées par le modèle logistique mixte appliqué à nos données.

9.1.3 SYNTHÈSE DU NOMBRE D'ERREURS DE LECTURE

Malgré le peu d'erreurs produites globalement, la précision varie en fonction de l'engagement moteur, du type d'amorce, de la congruence et de l'âge. En effet, l'exploration descriptive a montré un faible nombre d'erreurs de façon générale (4,48 % des syntagmes lus comportent des erreurs) ainsi que plusieurs tendances. Le nombre d'erreurs tend à être plus bas en condition -EM que +EM, plus bas en condition NL et L par rapport à M et plus bas en condition IC que C. Le nombre d'erreurs tend également à diminuer avec l'avancée de l'âge.

Les résultats du modèle binomial mixte inférentiel comportant uniquement les VI d'intérêt (Engagement moteur, Type d'amorce et Congruence Métrique) appliqué à nos données de production d'erreurs de lecture montrent un effet simple de l'Engagement moteur, modulé par le Type d'amorce. La production d'erreurs de lecture serait ainsi plus importante en condition +EM qu'en condition -EM, et ce pour les Type d'amorce L et NL, mais pas M. De plus, nous n'avons pas trouvé d'effet simple du Type d'amorce ou de la Congruence, ni des interactions « Engagement_moteur : Congruence » et « Type_amorce : Congruence ».

Ces résultats seront mis en regard du temps d'initiation de la parole, qui nous permettra de voir si l'augmentation du nombre d'erreurs peut être liée à une augmentation du temps d'initiation, tel que le suggère le paradigme d'amorçage (Shao & Meyer, 2018). Cependant, un temps de traitement raccourci peut également se faire au prix d'une diminution de la précision du traitement, et donc d'un nombre d'erreurs plus important (Levelt, 1999).

9.2 TEMPS D'INITIATION DE LA PAROLE

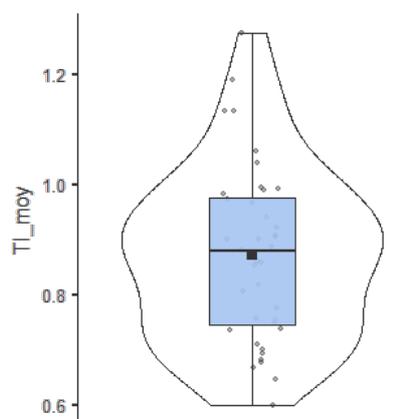
Pour rappel, le Temps d'Initiation (TI) de la parole correspond à l'intervalle temporel (silence) entre le moment d'apparition du syntagme cible à l'écran et le début de l'élocution du participant lisant ce syntagme à haute voix (cf. **Figure 27**, p. 119). Tout comme pour le nombre d'erreurs de lecture, le TI de la parole nous sert de mesure quant à l'observation de l'effet de différents amorçages rythmiques de la parole (nos variables indépendantes – VI) : l'Engagement Moteur du participant (EM), le Type d'Amorce (TA) et la Congruence Métrique (CM), afin de faire émerger la ou les conditions d'amorçage rythmique les plus facilitantes pour l'initiation de la parole.

9.2.1 EXPLORATION DESCRIPTIVE

Le Temps d'Initiation (TI) moyen de chaque participant a été agrégé à partir de la moyenne de l'ensemble des TI de chaque syntagme cible lu, après traitement des données aberrantes (cf. 6.3.2, p. 120). A partir de l'ensemble des TI moyens de chaque participant toutes conditions confondues ($n = 40$), nous avons pu calculer le TI global de notre échantillon. La distribution des TI moyens des individus suivant une distribution normale ($p = 0,393$ au test de normalité de Shapiro-Wilk, cf. **Figure 53** et **Tableau 23**), nous avons calculé la moyenne comme indice de tendance centrale. Pour l'ensemble de notre échantillon, nous avons ainsi trouvé un TI moyen global de 873 ms.

Figure 53 :

Représentation graphique du TI de la parole moyen des participants



Note. La boîte à moustaches indique la médiane et les quartiles 1 et 3 ; le carré la moyenne ; le violon la densité. TI_moy : TI de la parole moyen, en secondes.

Tableau 23 :*Valeurs descriptives du temps d'initiation de la parole moyen des participants*

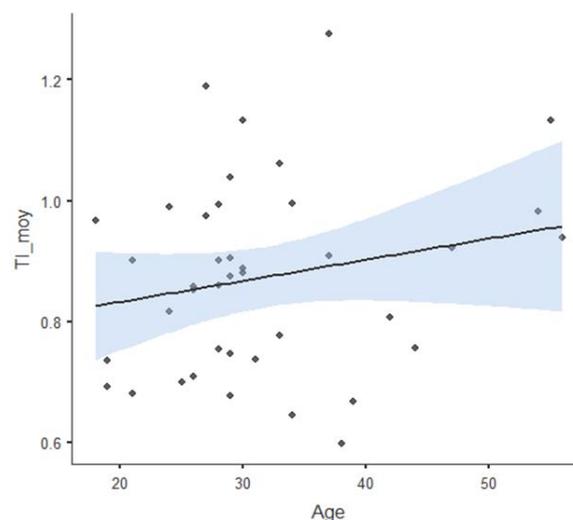
Statistiques descriptives

| | N | Moyenne | Médiane | Ecart-type | Minimum | Maximum | Shapiro-Wilk | |
|--------|----|---------|---------|------------|---------|---------|--------------|-------|
| | | | | | | | W | p |
| TI_moy | 40 | 0.873 | 0.878 | 0.158 | 0.599 | 1.28 | 0.971 | 0.393 |

Note. Une valeur $p < 0,05$ indique une distribution non paramétrique des données. TI_moy : TI de la parole moyen, en secondes.

De plus, en explorant les données (données agrégées), nous observons plusieurs tendances concernant le TI de la parole, selon les conditions expérimentales (EM, TA et CM). Les TI sont plus courts en condition d'amorçage avec engagement moteur (+EM) (moyenne = 856 ms) qu'en condition d'amorçage sans engagement moteur (-EM) (moyenne = 893 ms) ; les TI sont plus courts en condition d'amorçage vocal non langagier (NL) (moyenne = 860 ms) qu'en condition d'amorçage musical (M) (moyenne = 874 ms) et qu'en condition d'amorçage langagier (L) (moyenne = 890 ms) ; et les TI sont de durée équivalente en conditions incongruente (IC) (moyenne = 873 ms) et congruente (C) (moyenne = 876 ms).

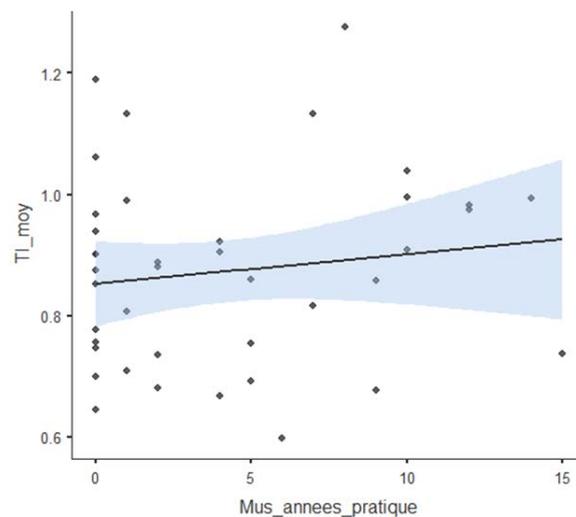
Nous avons également observé l'effet des caractéristiques socio-démographiques sur le TI. Nous rappelons cependant que ces caractéristiques n'ont pas été contrôlées spécifiquement pour cette étude et sont donc à interpréter avec précaution. Il semblerait que les TI s'allongent avec l'avancée de l'âge et l'augmentation du nombre d'années de pratique de la musique (cf. **Figure 54** et **Figure 55**).

Figure 54 :*Evolution du TI de la parole en fonction de l'âge*

Note. La ligne indique la droite de régression linéaire et l'espace coloré l'erreur standard. TI_moy : TI de la parole moyen, en secondes.

Figure 55 :

Evolution du TI de la parole en fonction du nombre d'années de musique



Note. La ligne indique la droite de régression linéaire et l'espace coloré l'erreur standard. TI_moy : TI de la parole moyen, en secondes.

De plus, les TI sont plus courts pour les participants joueurs de jeux vidéo (médiane = 855 ms) que pour les non-joueurs (médiane = 909 ms).

9.2.2 ANALYSE INFÉRENTIELLE

Afin de répondre à nos *Hypothèse 2 à 5* et de modéliser la relation entre les différences d'amorçage rythmique et le Temps d'Initiation (TI) de la parole, nous avons ensuite appliqué un modèle linéaire mixte à nos données ($n = 3654^{105}$), avec en effet fixe les différences d'amorçage rythmique (Engagement Moteur, Type d'Amorce et Congruence Métrique), en effet aléatoire les participants et les cibles et en variable dépendante le TI de la parole.

Dans un premier temps, nous avons appliqué à nos données un modèle comprenant uniquement nos variables indépendantes d'intérêt (EM, TA et CM) (cf. 9.2.2.1). Dans un deuxième temps, nous avons comparé ce premier modèle avec un modèle incluant également des variables contrôles socio-démographiques afin d'estimer la potentielle interaction entre les caractéristiques socio-démographiques et les effets d'intérêt (cf. 9.2.2.2, p. 181). Enfin, dans un

¹⁰⁵ Données réelles/désagrégées : 40 participants * 12 conditions expérimentales * 8 observations, sans les valeurs manquantes.

troisième modèle, nous avons comparé l'effet des variables d'intérêt (EM, TA et CM) avec le groupe métrique de l'amorce (cf. 9.2.2.3, p. 182).

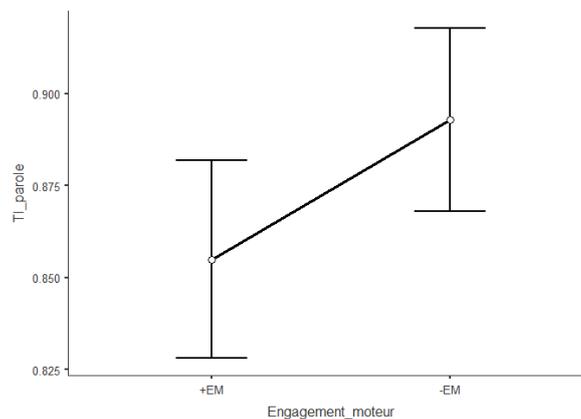
9.2.2.1 EFFET DES DIFFERENTES CONDITIONS D'AMORÇAGE RYTHMIQUE SUR LE TI DE LA PAROLE

Après avoir établi la structure aléatoire la plus ajustée à notre modèle contenant uniquement les variables d'intérêt (EM, TA et CM) (cf. ANNEXES – PARTIE 2, document complémentaire, 1.2.2, p. 76), nous avons appliqué celui-ci à nos données de TI de la parole.

Nous observons un effet principal de l'Engagement moteur ($Estimate = 0,038$; $p = 0,007$) sur le TI de la parole. Le TI de la parole est plus court en situation d'amorçage avec Engagement moteur (+EM) que sans Engagement moteur (-EM) (cf. **Figure 56**).

Figure 56 :

Comparaison du TI de la parole selon l'Engagement Moteur des participants

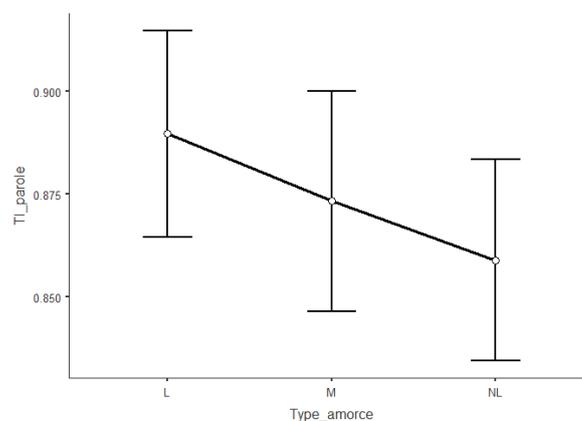


Note. Comparaison du TI de la parole (TI_parole, en secondes) en fonction de l'Engagement Moteur des participants (« +EM » = avec EM ; « -EM » = sans EM). La droite représente l'estimation (différence des moyennes estimées) de la modélisation et les barres l'erreur standard.

Ce résultat concernant l'effet de l'engagement moteur est cohérent avec l'*Hypothèse 2*, à propos de l'existence d'un partage anatomo-fonctionnel entre la motricité manuelle et les gestes de production de la parole, et selon laquelle l'engagement moteur manuel stimule les gestes de production de la parole et facilite ainsi l'initiation de la parole.

Nous observons également un effet principal du Type d'amorce ($p = 0,001$) sur le TI de la parole. Les comparaisons post-hoc (correction de Bonferroni) montrent une différence significative entre le type d'amorce non langagier (NL) et le type d'amorce langagier (L) ($Difference = 0,031$; $p = 0,001$) tandis que les autres différences de types d'amorces ne sont pas significatives (cf. Figure 57).

Figure 57 :
Comparaison du TI de la parole en fonction du Type d'Amorce



Note. Comparaison du TI de la parole (TI parole, en secondes) en fonction du Type d'amorce (« L » = langagier ; « M » = musical et « NL » = vocal non langagier). Les droites représentent l'estimation (différences des moyennes estimées) de la modélisation et les barres l'erreur standard.

Ce résultat concernant le type d'amorce est partiellement cohérent avec l'*Hypothèse 3*, selon laquelle un amorçage langagier (L) serait plus facilitant qu'un amorçage non langagier (NL), ce dernier étant plus facilitant qu'un amorçage musical (M). Dans nos résultats, la tendance (effet non significatif) d'un effet facilitant pour le TI de la parole de l'amorce NL par rapport à l'amorce M est cohérente avec l'*Hypothèse 3.2* selon laquelle un amorçage vocal (L ou NL) est plus facilitant qu'un amorçage non vocal musical (M) en raison de l'activation du système moteur de production de la parole lors de la perception de parole (amorce L ou NL) qui favorise la production cible subséquente. En revanche, concernant l'*Hypothèse 3.1*, nos résultats montrent un effet plus facilitant (effet significatif) de NL par rapport à L et une tendance plus facilitante de M par rapport à L (effet non significatif), qui pourrait indiquer que, puisque ces deux amorces (NL et M) sont neutres d'un point de vue sémantique et syntaxique, elle ne créent ainsi pas de contexte préalable à la lecture des syntagmes cibles qui viendrait interférer avec l'initiation de la parole, à la différence d'une amorce langagière.

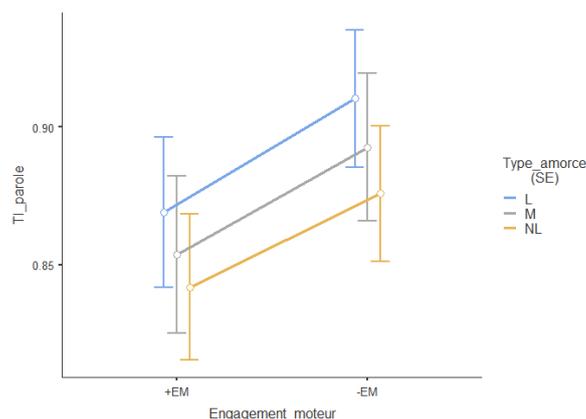
Nous n'observons pas d'effet principal de la congruence ($p = 0,591$) ni d'effet en interaction avec l'engagement moteur ou le type d'amorce. Ce résultat vient contredire l'*Hypothèse 4*, selon laquelle nous nous attendions à observer un effet facilitant d'un amorçage métriquement congruent par rapport à un amorçage métriquement incongruent, ainsi que l'*Hypothèse 5*, selon laquelle l'effet de la congruence métrique est renforcé par l'engagement moteur du participant. Ces résultats pourraient cependant s'expliquer d'une part par le fait que nous n'avons pas inclus dans notre protocole autant de granularité dans les conditions de

congruence que l'étude de Cason, Astésano, et al. (2015), par souci de faisabilité afin de ne pas surcharger le nombre de conditions expérimentales. D'autre part, ces résultats pourraient également être le reflet d'une interaction entre le groupe métrique de l'amorce et l'effet de la congruence métrique. Nous reviendrons sur cette interaction dans un nouveau modèle dans la partie 9.2.2.3 (p. 182).

En revanche, nous observons un effet cumulatif de l'engagement moteur et du type d'amorce : l'effet facilitant de l'amorçage +EM se retrouve pour tous les types d'amorce de manière significative (L : $p = 0,008$; NL : $p = 0,026$; M : $p = 0,011$) (cf. **Figure 58**). Ce résultat semble ainsi cohérent avec l'hypothèse OPERA (Patel, 2011), qui postule l'existence de mécanismes cognitifs partagés pour le traitement de la musique et du langage, ainsi qu'avec le caractère supramodal du rythme.

Figure 58 :

Comparaison du TI de la parole en fonction de EM et TA



Note. Comparaison du TI de la parole (TI parole, en secondes) en fonction de l'Engagement Moteur (« +EM » et « -EM ») du participant et du Type d'Amorce (« L » = langagier ; « M » = musical et « NL » = vocal non langagier). Les droites représentent l'estimation (différences des moyennes estimées) de la modélisation et les barres l'erreur standard.

A partir des moyennes marginales estimées par le modèle (cf. **Tableau 24**), l'amorçage rythmique le plus facilitant pour l'initiation de la parole (TI le plus court) est l'amorçage avec engagement moteur non langagier (+EM_NL = 842 ms) et le moins facilitant est l'amorçage sans engagement moteur langagier (-EM_L = 910 ms). Ces résultats pourraient servir de piste pour l'élaboration de protocoles de remédiation langagière, notamment pour les pathologies présentant un trouble de l'initiation de la parole, telles que le bégaiement ou l'aphasie non fluente.

Tableau 24 :*Moyennes marginales estimées du TI de la parole, selon EM et TA*

Estimate Marginal Means - Engagement_moteur * Type_amorce

| Engagement_moteur | Type_amorce | Mean | SE | df | 95% Confidence Intervals | |
|-------------------|-------------|-------|--------|------|--------------------------|-------|
| | | | | | Lower | Upper |
| -EM | L | 0.910 | 0.0250 | 41.1 | 0.860 | 0.961 |
| -EM | M | 0.893 | 0.0269 | 40.7 | 0.838 | 0.947 |
| -EM | NL | 0.876 | 0.0246 | 41.1 | 0.826 | 0.925 |
| +EM | L | 0.869 | 0.0273 | 40.7 | 0.814 | 0.924 |
| +EM | M | 0.854 | 0.0285 | 40.5 | 0.796 | 0.911 |
| +EM | NL | 0.842 | 0.0266 | 40.8 | 0.788 | 0.896 |

Note. Moyennes marginales du TI de la parole estimées par le modèle linéaire mixte appliqué à nos données (colonne Mean, en secondes).

9.2.2.2 EFFET DES DIFFERENTES CONDITIONS D'AMORÇAGE RYTHMIQUE ET DES CARACTERISTIQUES SOCIO-DEMOGRAPHIQUES SUR LE TI DE LA PAROLE

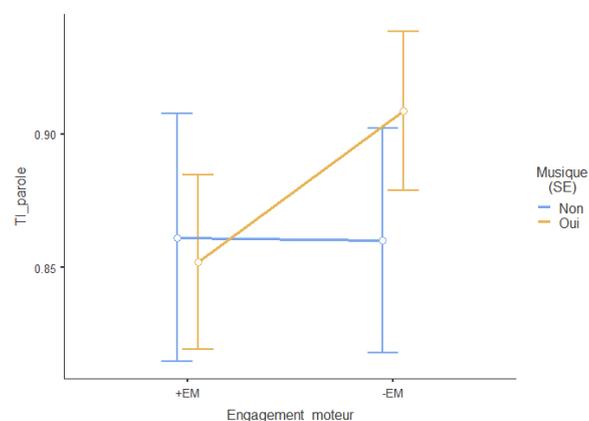
Après avoir estimé l'effet de l'engagement moteur, du type d'amorce et de la congruence métrique (variables d'intérêt) sur le Temps d'Initiation (TI) de la parole, nous avons voulu aller plus loin en ajoutant au modèle des variables socio-démographiques (variables contrôles). Ces variables contrôles permettent de contrôler si la variance observée est liée à une interaction entre les variables indépendantes d'intérêt et d'éventuelles variables contrôles.

Nous sommes donc partie d'un modèle saturé avec les paramètres suivants : facteurs d'intérêt « Engagement_moteur », « Type_amorce » et « Congruence » ; covariables contrôles « Age », « Mus_annees_pratique », « Danse_annees_pratique » ; facteurs contrôles « Sexe », « Musique », « Danse » et « Jeux_video » ; nous avons inclus les interactions doubles entre les facteurs d'intérêt et entre d'une part les facteurs d'intérêt et d'autre part les paramètres contrôles. Nous n'avons pas inclus les interactions doubles entre paramètres contrôles qui auraient surchargé le modèle. Après avoir comparé pas à pas la pertinence des paramètres pour parvenir à un modèle le plus simple et le plus ajusté (cf. ANNEXES – PARTIE 2, document complémentaire, 1.2.2, p. 84), nous avons appliqué le nouveau modèle à nos données.

Nous observons que l'effet de l'Engagement Moteur (EM) des participants sur le TI de la parole est modulé par la pratique de la musique ($p = 0,041$). Chez les musiciens, le TI de la parole est plus court en situation d'amorçage +EM que d'amorçage -EM ($Estimate = 0,057$; $p < 0,001$), tandis que l'effet de l'Engagement moteur n'est pas significatif chez les non-musiciens ($p = 0,966$) (cf. **Figure 59**). Nous rappelons cependant que les caractéristiques socio-démographiques non pas été contrôlées de manière spécifique pour cette étude et que ce résultat est de fait à interpréter avec précaution. De plus, il est important de noter que nous avons considéré un participant comme « musicien » dans cette étude dès lors qu'il a indiqué pratiquer ou avoir pratiqué une activité musicale.

Figure 59 :

Comparaison du TI de la parole selon l'Engagement Moteur et la pratique de la musique



Note. Comparaison du TI de la parole (TI_parole, en secondes) en fonction de l'Engagement Moteur des participants (« +EM » = avec EM ; « -EM » = sans EM) et de la pratique de la musique (Oui/Non). Les droites représentent l'estimation (différences des moyennes estimées) de la modélisation et les barres l'erreur standard.

Nous n'observons pas d'autre interaction des effets d'intérêt (Engagement Moteur, Type d'Amorce ou Congruence Métrique) avec les caractéristiques socio-démographiques.

9.2.2.3 EFFET DES DIFFERENTES CONDITIONS D'AMORÇAGE RYTHMIQUE ET DU GROUPE METRIQUE DE L'AMORCE SUR LE TI DE LA PAROLE

Après avoir estimé l'effet de l'engagement moteur, du type d'amorce et de la congruence métrique (variables d'intérêt) sur le Temps d'Initiation (TI) de la parole, ainsi que les possibles interactions avec les caractéristiques socio-démographiques des participants, nous avons voulu aller plus loin en ajoutant au modèle la variable contrôle du Groupe Métrique (GM)¹⁰⁶ de l'amorce. Cela nous permet de contrôler si la variance observée dans le TI est liée à une

¹⁰⁶ Pour un rappel de nos quatre Groupes Métriques, se référer au **Tableau 3**, p. 94.

interaction entre les variables indépendantes d'intérêt et cette éventuelle variable contrôle. En effet, nous réutilisons une partie du matériel expérimental de l'étude de Cason, Astésano, et al. (2015) qui rapportent une interaction entre l'effet de Congruence Métrique (CM) et le Groupe Métrique (GM) de l'amorce, l'effet de CM n'étant significatif dans leur étude que pour les amorces de structure métrique « x x X x x X » ou « x x X x x x X »¹⁰⁷ mais pas pour les amorces de structure métrique « x x x X x X » ou « x x x X x x X »¹⁰⁸.

Nous sommes donc partie d'un modèle saturé avec les paramètres suivants : facteurs d'intérêt « Engagement_moteur », « Type_amorce » et « Congruence » ; covariable contrôle « Groupe_amorce » et toutes les interactions doubles impliquant les facteurs d'intérêt. Après avoir comparé pas à pas la pertinence des paramètres pour parvenir à un modèle le plus simple et le plus ajusté (cf. ANNEXES – PARTIE 2, document complémentaire, 1.2.2, p. 110), nous avons appliqué le nouveau modèle à nos données.

Tout comme pour le modèle comportant uniquement les VI d'intérêt (EM, TA et CM) (9.2.2.1, p. 178), nous retrouvons un effet de l'Engagement Moteur ($p = 0,008$) et du Type d'Amorce ($p = 0,001$). A nouveau, les comparaisons post-hoc (correction de Bonferroni) montrent une différence significative entre le Type d'Amorce non langagier (NL) et le type d'amorce langagier (L) ($Difference = 0,031$; $p = 0,001$) tandis que les autres différences de types d'amorces ne sont pas significatives. Nous retrouvons également l'effet cumulatif de l'Engagement Moteur et du Type d'Amorce : l'effet facilitant de l'amorçage +EM se retrouve pour tous les Types d'Amorce de manière significative (L : $p = 0,008$; NL : $p = 0,026$; M : $p = 0,011$).

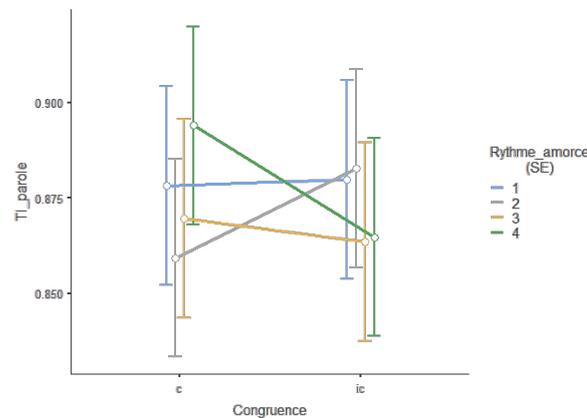
Enfin, nous observons une interaction de la Congruence Métrique (CM) avec le Groupe Métrique de l'amorce ($p = 0,035$), l'effet de CM n'étant retrouvé que pour les amorces des Groupes Métriques 2 (x x X x x X) ($Estimate = 0,024$; $p = 0,033$) et 4 (x x x X x x X) ($Estimate = -0,029$; $p = 0,008$) mais pas pour les Groupes Métriques 1 (x x x X x X) ($p = 0,884$) et 3 (x x X x x x X) ($p = 0,579$). L'amorçage métriquement congruent est donc facilitant uniquement pour les amorces du GM 2, et l'amorçage métriquement incongruent est étonnamment facilitant pour les amorces du GM 4. Nous retrouvons ici partiellement les résultats de Cason, Astésano, et al. (2015), qui avaient montré un effet significatif pour les amorces des groupes métriques 2 (x x X x x X) et 3 (x x X x x x X).

¹⁰⁷ Dans notre étude, ces structures correspondent aux GM 2 et 3 (cf. 3.2.3.1, p. 93)

¹⁰⁸ Dans notre étude, ces structures correspondent aux GM 1 et 4 (cf. 3.2.3.1, p. 93)

Figure 60 :

Comparaison du TI de la parole en fonction de CM et du groupe métrique de l'amorce



Note. Comparaison du TI de la parole (TI_parole, en secondes) en fonction de la Congruence Métrique (C, IC) et du Groupe Métrique de l'amorce (Rythme_amorce). Les droites représentent l'estimation (différences des moyennes estimées) de la modélisation et les barres l'erreur standard.

9.2.3 SYNTHÈSE DU TI DE LA PAROLE

L'exploration descriptive du TI de la parole montre une tendance centrale pour le TI global (moyenne) de 873 ms ainsi que plusieurs tendances. Les TI tendent à être plus courts en condition d'amorçage +EM que -EM, plus courts en condition NL que M et les plus longs pour L, et de durée équivalente en conditions IC et C. Les TI tendent à s'allonger avec l'avancée de l'âge et du nombre d'années de pratique de la musique. Les TI tendent à être plus courts pour les participants joueurs de jeux vidéo.

Les résultats du premier modèle inférentiel comportant uniquement nos variables explicatives d'intérêt « Engagement_moteur », « Type_amorce » et « Congruence » montrent un effet simple de l'Engagement Moteur (EM) et du Type d'Amorce (TA) sur le TI de la parole, mais pas de la Congruence Métrique (CM) ou de l'interaction entre EM et CM. Le TI de la parole est plus court en situation d'amorçage rythmique +EM que -EM, et plus court en situation d'amorçage rythmique NL que L. De plus, l'effet facilitant de +EM par rapport à -EM se retrouve pour chacun des types d'amorce (NL, L et M). À partir des moyennes marginales estimées par le modèle, l'amorçage rythmique le plus facilitant pour l'initiation de la parole (TI le plus court) est l'amorçage avec engagement moteur non langagier (+EM_NL = 842 ms) tandis que le moins facilitant est l'amorçage sans engagement moteur langagier (-EM_L = 910 ms).

L'effet significatif de l'Engagement Moteur (EM) est cohérent avec l'*Hypothèse 2*, selon laquelle l'engagement moteur manuel du participant (+EM) stimule les gestes de production de la parole et facilite l'initiation de la parole.

L'effet du Type d'Amorce (TA) est partiellement cohérent avec l'*Hypothèse 3*. Selon l'*Hypothèse 3.2*, nous nous attendions à un effet facilitant de l'amorce L sur l'amorce NL en raison de l'effet de lexicalité de l'amorce L. Cependant, nos résultats montrent l'effet inverse : l'amorce NL est plus facilitante pour la l'initiation de la parole que l'amorce L, ce qui pourrait être interprété comme un avantage d'une amorce neutre d'un point de vue syntaxique et sémantique, à la différence de l'amorce L, dont le contenu lexical viendrait interférer avec le traitement de la cible lexicale subséquente. Selon l'*Hypothèse 3.1*, nous nous attendions à un effet facilitant d'une amorce rythmique vocale (L ou NL) sur une amorce rythmique non vocale musicale (M). Cet effet est partiellement retrouvé, concernant la différence (non significative) entre l'amorce NL et l'amorce M. La tendance d'un effet facilitant de l'amorce NL (vocale) sur l'amorce M (non vocale) suggère l'activation du système moteur de production de la parole lors de l'écoute de l'amorce NL, facilitant l'initiation de parole subséquente. Nous avons trouvé un effet facilitant significatif de NL par rapport à M sur le TI de la parole dans de précédentes analyses (Gindre et al., 2023), qui diffèrent de celles-ci en ce que nous ici adopté une approche plus conservatrice. En effet, dans ce présent travail, nous avons exclu de l'analyse les données d'un participant dont le profil se différenciait du reste de l'échantillon et avons appliqué une approche de modélisation linéaire plus stricte basée sur la comparaison et la sélection des effets aléatoires et des effets fixes.

L'absence d'effet simple de la congruence métrique a pu être expliqué en raison d'une interaction entre l'effet de congruence métrique et l'effet du groupe métrique de l'amorce. Nous avons ainsi trouvé un effet facilitant de l'amorçage congruent pour une amorce du groupe métrique 2 (x x X x x X).

En ajustant mieux notre modèle à notre échantillon (ajout de variables explicatives contrôles représentatives des caractéristiques socio-démographiques), nous retrouvons un effet de l'Engagement Moteur (modulé par la pratique de la musique) et du Type d'Amorce sur le TI de la parole, et pas d'effet de la Congruence Métrique (simple ou en interaction). Nous observons que l'effet facilitant de +EM par rapport à -EM sur le TI de la parole n'est présent que chez les participants musiciens. Cependant, en raison du caractère non contrôlé des

variables socio-démographiques dans notre étude, l'effet de la pratique musicale est néanmoins à interpréter avec précaution.

La comparaison des résultats d'analyses statistiques issues de ces trois modèles nous invite à prendre en considération et mieux contrôler pour de futures recherches la pratique de la musique ainsi que le groupe métrique des amorces et les niveaux de granularité de la Congruence Métrique, qui viennent nuancer les résultats du premier modèle.

10 ACTIVITES 1 ET 2 : RESULTATS CROISES

Dans cette dernière partie des résultats, nous présentons les analyses croisées entre les différentes tâches des deux activités, afin d'investiguer l'Hypothèse 1, à propos de l'existence d'une période de référence supramodale dans une zone allant de 400 à 800 ms (avec une moyenne aux alentours de 600 ms) et l'Hypothèse 6, à propos de l'existence de mécanismes partagés pour le traitement du rythme et du langage.

Dans un premier temps (cf. 10.1), nous mettons en regard les mesures de Tempo Moteur Spontané (TMS) moyen, tempo de la Synchronisation Sensori-Motrice (SMS) le plus stable et le plus précis (Activité 1) et de Temps d'Initiation (TI) de la parole.

Dans un second temps (cf. 10.2, p. 191), nous mettons en regard les capacités de traitement du rythme (stabilité et précision du TMS et de la SMS) et les capacités de traitement du langage (nombre d'erreurs de lecture et TI de la parole).

10.1 PERIODE DE REFERENCE SUPRAMODALE

10.1.1 ACTIVITE 1

Concernant l'Activité 1, nous avons émis l'Hypothèse 1.1 selon laquelle la Synchronisation Sensori-Motrice (SMS) est la plus stable et la plus précise au tempo (parmi les trois proposés : 225, 600 et 900 ms) se rapprochant le plus du Tempo Moteur Spontané (TMS) moyen.

Dans la partie 8.1 (p. 134), nous avons rapporté un TMS moyen global de l'échantillon de 712 ms, qui se situe bien dans une zone entre 400 et 800 ms. Dans la partie 8.2 (p. 145), nous avons observé que les participants ont eu plus de mal à mettre en place la Synchronisation Sensori-Motrice (SMS) au tempo de métronome d'Intervalle Inter-Stimuli (ISI) 225 ms par rapport aux tempi 600 et 900 ms, cette plus grande difficulté au tempo 225 ms étant reflétée par :

- un nombre plus important de séquences ne vérifiant pas le test de Rayleigh¹⁰⁹ pour le tempo de SMS 225 ms (92,31 % des séquences non synchronisées), par rapport aux tempi 600 ms (7,7 %) et 900 ms (0 %) (cf. **Tableau 25**) ;

Tableau 25 :

Tableau de fréquences des vecteurs circulaires non significatifs au test de Rayleigh

| Tempi_SMS | Quantités | % du Total |
|-----------|-----------|------------|
| 225_ms | 12 | 92,31 % |
| 600_ms | 1 | 7,7 % |
| 900_ms | 0 | 0 % |

Note. Un vecteur non significatif au test de Rayleigh témoigne d'une non réussite de la mise en place de la SMS, en fonction du tempo.

- un nombre plus important de séquences présentant une asynchronie positive¹¹⁰ pour le tempo de SMS 225 ms (74,34 %), par rapport aux tempi 600 ms (10,26 %) et 900 ms (15,38 %) (cf. **Tableau 26**).

Tableau 26 :

Tableau de fréquences de l'asynchronie positive de la SMS, en fonction du tempo.

| Tempi_SMS | Quantités | % du Total |
|-----------|-----------|------------|
| 225_ms | 29 | 74,34 % |
| 600_ms | 4 | 10,26 % |
| 900_ms | 6 | 15,38 % |

Note. Une asynchronie positive témoigne d'une réaction aux battements de métronome plutôt que d'une anticipation.

- une différence significative de stabilité de la SMS entre d'une part le tempo 225 ms (moins stable) et d'autre part les tempi 600 et 900 ms (plus stables) (cf. 8.2.1.2.1, p. 151)
- une tendance à une précision de la SMS plus faible (asynchronie plus élevée) pour le tempo 225 ms, suivi par le tempo 600 ms et enfin 900 ms, bien que ces différences ne soient pas significatives (8.2.2.1.1, p. 156).

De plus, à partir des tempi auxquels la SMS est la plus stable et la plus précise (cf. 8.2.3, p. 165), nous avons observé que :

¹⁰⁹ Pour rappel, le test de Rayleigh permet de distinguer la répartition des frappes du participant d'une répartition aléatoire, et témoigne de la réussite de la mise en place de la synchronisation.

¹¹⁰ Pour rappel, une asynchronie positive témoigne d'une réaction aux battements de métronome, à la différence d'une asynchronie négative indiquant une anticipation et donc la mise en place de la synchronisation.

- les participants présentant une SMS la plus stable au tempo de 900 ms (par rapport aux tempi plus rapides 225 ms et 600 ms) ont tendance à présenter un TMS moyen plus lent (cf. **Tableau 27**), cette différence n'étant cependant pas significative (test $t : p = 0,559$)

Tableau 27 :

Valeurs descriptives du TMS moyen en fonction du tempo le plus stable de la SMS

Statistiques descriptives

| | Stabilité_max | N | Moyenne | Médiane | Ecart-type | Minimum | Maximum | Shapiro-Wilk | |
|---------|---------------|----|---------|---------|------------|---------|---------|--------------|-------|
| | | | | | | | | W | p |
| TMS_Moy | 600_ms | 13 | 0.738 | 0.689 | 0.198 | 0.450 | 1.11 | 0.951 | 0.610 |
| | 900_ms | 27 | 0.785 | 0.736 | 0.252 | 0.451 | 1.48 | 0.909 | 0.022 |

Note. Valeur du TMS moyen (en secondes) en fonction du tempo le plus stable de la SMS. Aucun participant n'a présenté une synchronisation la plus stable pour un tempo de 225 ms.

- les participants présentant une SMS la plus précise au tempo de 900 ms (par rapport aux tempi plus rapides 225 ms et 600 ms) ont tendance à présenter un TMS moyen plus lent (cf. **Tableau 28**), cette différence n'étant pas significative (test $t : p = 0,480$).
- Au demeurant, la variabilité inter-individuelle est plus réduite pour le TMS moyen des participants dont le tempo de la SMS le plus stable est à 600 ms (la valeur de l'écart-type est de 190 ms, inférieure aux deux autres valeurs pour 225 ms et 900 ms).

Tableau 28 :

Valeurs descriptives du TMS moyen en fonction du tempo le plus précis de la SMS.

Statistiques descriptives

| | Précision_max | N | Moyenne | Médiane | Ecart-type | Minimum | Maximum | Shapiro-Wilk | |
|---------|---------------|----|---------|---------|------------|---------|---------|--------------|-------|
| | | | | | | | | W | p |
| TMS_Moy | 225_ms | 13 | 0.756 | 0.722 | 0.221 | 0.450 | 1.12 | 0.929 | 0.328 |
| | 600_ms | 13 | 0.720 | 0.689 | 0.190 | 0.451 | 1.11 | 0.938 | 0.434 |
| | 900_ms | 14 | 0.829 | 0.771 | 0.283 | 0.481 | 1.48 | 0.885 | 0.068 |

Note. Valeur du TMS moyen (en secondes) en fonction du tempo le plus précis de la SMS

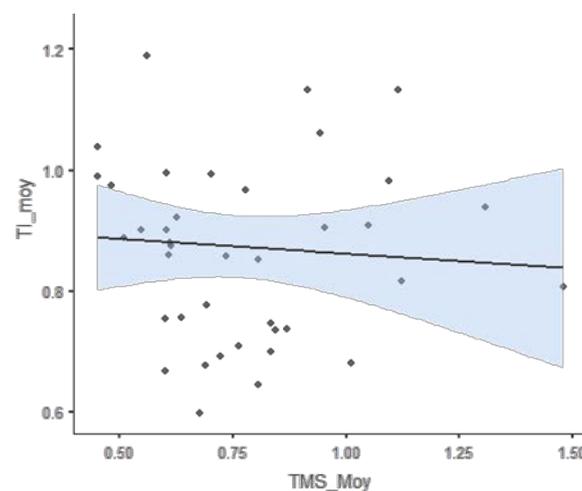
Les participants ont ainsi montré plus de facilité pour mettre en place la SMS à un tempo se rapprochant du TMS moyen global de 712 ms, tempi de SMS 600 et 900 ms, par rapport au tempo de SMS le plus éloigné, 225 ms. De plus, les participants présentant une période de référence plus lente (TMS moyen) sont ceux qui ont présenté un tempo le plus stable et le plus précis le plus lent (900 ms). Ces résultats sont ainsi cohérents avec l'*Hypothèse 1.1*.

10.1.2 ACTIVITES 1 ET 2

Concernant à présent l'*Hypothèse 1.2*, selon laquelle le Temps d'Initiation (TI) de la parole représente la période de référence vocale et est corrélé au TMS moyen, nous avons observé un TI moyen global de l'échantillon de 873 ms. Ce résultat est donc un peu plus lent que les fréquences situées dans la zone s'étendant de 400 à 800 ms pour la période de référence du TMS (Drake, 2010), et un peu plus lent que les TI habituellement recensés dans la littérature (cf. 1.2.1.3.1, p. 19). Cette différence pourrait s'expliquer par le fait que le TI est généralement mesuré dans des tâches de lecture de mots isolés ou de dénomination d'image (mot), tandis que nous avons ici choisi une tâche linguistiquement plus complexe, ce qui pourrait se manifester par un allongement du temps d'initiation de la parole. Nous n'observons pas de corrélation entre le TMS moyen et le TI moyen (r de Spearman = -0,149 ; $p = 0,356$) (cf. **Figure 61**).

Figure 61 :

Evolution du TI de la parole moyen en fonction du TMS moyen



Note. Droite de régression illustrant la relation entre le TMS moyen (TMS_moy, en secondes) et le TI de la parole moyen (TI_moy, en secondes) des participants ($n = 40$). L'espace coloré représente l'erreur standard.

10.2 LIENS ENTRE LES MECANISMES DE TRAITEMENT DU RYTHME ET DU LANGAGE

Afin d'investiguer l'Hypothèse 6, à propos de l'existence de mécanismes partagés pour le traitement du rythme et du langage, nous avons recueilli différentes mesures représentatives des capacités de traitement du rythme (la stabilité du Tempo Moteur Spontané, la stabilité et la précision de la Synchronisation Sensori-Motrice) dans l'Activité 1 et des capacités de traitement du langage (nombre d'erreurs de lecture et TI de la parole) dans l'Activité 2, que nous avons comparées entre elles à l'aide de corrélations de rang de Spearman (cf. **Tableau 29**, p. 193).

Concernant tout d'abord les capacités de traitement du rythme (Activité 1), nous n'avons trouvé aucune corrélation entre d'une part le TMS moyen et d'autre part les mesures de stabilité (TMS et SMS) et de précision (SMS) rythmiques, la vitesse du TMS moyen ne semble donc pas liée aux capacités de traitement du rythme.

En revanche, les différentes mesures de stabilité et de précision du TMS et de la SMS présentent de nombreuses corrélations significatives les unes avec les autres. Nous pouvons notamment relever que la stabilité du TMS et la stabilité de la SMS au tempo 225 ms apparaissent être les deux mesures les plus représentatives des capacités de traitement du rythme, en étant significativement corrélées avec l'ensemble des mesures de stabilité et de précision du TMS et de la SMS. Les individus les plus stables lors de la tâche de TMS ou de SMS au tempo 225 ms sont ainsi les plus stables et les plus précis pour toutes les autres mesures des capacités de traitement du rythme. Ces deux mesures pourraient ainsi être les mesures de référence dans la recherche des liens entre les capacités de traitement du rythme et du langage.

En second lieu, concernant le lien entre les capacités de traitement du rythme et du langage, nous observons une corrélation entre le Temps d'Initiation (TI) de la parole et la stabilité de la SMS au tempo 600 ms ($p < 0,05$), indiquant que les participants les plus stables lors de la tâche de SMS au tempo de 600 ms sont ceux qui présentent un TI de la parole le plus lent. Nous n'observons pas d'autre corrélation significative, que ce soit entre le TI moyen et le nombre d'erreurs (Activité 2) ou bien entre d'une part les capacités de traitement du langage et d'autre part les capacités de traitement du rythme.

Nous pouvons cependant noter plusieurs tendances :

- Les participants dont le Temps d'Initiation de la parole est le plus lent sont ceux qui produisent le moins d'erreurs ;
- Les participants les plus stables (valeur \bar{R} élevée) et les plus précis (valeur $\bar{\theta}$ basse) dans la tâche de SMS (tous tempi confondus), et ce notamment au tempo 225 ms, sont également ceux qui présentent le moins d'erreurs de lecture ;
- Les participants les plus stables et les plus précis dans la tâche de SMS (tous tempi confondus), sont également ceux qui présentent un TI de la parole le plus lent.

Ces résultats semblent indiquer d'une part qu'un contrôle accru lors de l'Activité 2 (lecture après amorçage rythmique) se traduit par un temps d'initiation de la parole plus lent, ce qui permet de réduire le nombre d'erreurs produites. D'autre part, ces résultats semblent indiquer que les participants les plus stables et les plus précis sont également ceux qui produisent le moins d'erreurs de lecture, ce qui invite à poursuivre la recherche sur les mécanismes de traitements partagés entre le rythme et le langage, notamment dans une perspective d'intégrer l'évaluation des capacités de traitement du rythme en lien avec l'évaluation des capacités de traitement du langage en contexte de remédiation.

Tableau 29 :
Matrice de corrélation entre les mesures de l'Activité 1 et l'Activité 2

Matrice de corrélation

| | TMS_Moy | Stabilité_ TMS | Stabilité_ 225_ms | Stabilité_ 600_ms | Stabilité_ 900_ms | moyenne_ Stabilité_ tous_tempi | Précision_ 225_ms | Précision_ 600_ms | Précision_ 900_ms | moyenne_ Précision_ tous_tempi | TI_moy | nb_ erreurs |
|---------------------------------|---------|----------------|-------------------|-------------------|-------------------|--------------------------------|-------------------|-------------------|-------------------|--------------------------------|--------|-------------|
| TMS_Moy | — | | | | | | | | | | | |
| Stabilité_TMS | -0.208 | — | | | | | | | | | | |
| Stabilité_225_ms | -0.165 | 0.534 *** | — | | | | | | | | | |
| Stabilité_600_ms | -0.165 | 0.503 ** | 0.598 *** | — | | | | | | | | |
| Stabilité_900_ms | -0.057 | 0.571 *** | 0.660 *** | 0.762 *** | — | | | | | | | |
| moyenne_ Stabilité_ tous_ tempi | -0.112 | 0.575 *** | 0.971 *** | 0.716 *** | 0.761 *** | — | | | | | | |
| Précision_225_ms | 0.124 | -0.354 * | -0.733 *** | -0.320 | -0.284 | -0.675 *** | — | | | | | |
| Précision_600_ms | 0.217 | -0.533 *** | -0.550 *** | -0.833 *** | -0.692 *** | -0.663 *** | 0.275 | — | | | | |
| Précision_900_ms | 0.153 | -0.511 *** | -0.592 *** | -0.717 *** | -0.723 *** | -0.699 *** | 0.319 | 0.817 *** | — | | | |
| moyenne_ précision_ tous_ tempi | 0.183 | -0.560 *** | -0.831 *** | -0.720 *** | -0.612 *** | -0.870 *** | 0.775 *** | 0.769 *** | 0.783 *** | — | | |
| TI_moy | -0.149 | -0.041 | 0.237 | 0.323 * | 0.213 | 0.232 | -0.135 | -0.191 | 0.009 | -0.127 | — | |
| nb_erreurs | 0.026 | 0.152 | -0.262 | -0.140 | 0.024 | -0.206 | 0.191 | 0.097 | 0.045 | 0.218 | -0.152 | — |

Note. * p < .05, ** p < .01, *** p < .001

Note. Matrice de corrélation entre les mesures de stabilité et précision du TMS, de la SMS (Activité 1) et du TI de la parole moyen et du nombre d'erreurs (Activité 2) (corrélations de Spearman).

V. DISCUSSION ET PERSPECTIVES

Dans cette étude, nous nous sommes questionnée sur les liens entre traitements du rythme et du langage, question initiée par un intérêt concernant la comparaison entre les traitements impliqués pour la musique et le langage ainsi que l'utilisation de la musique comme moyen thérapeutique en cas de trouble du langage. Plus précisément, nous avons investigué le partage des mécanismes de traitement du rythme et du langage au moyen de deux activités expérimentales, l'une renseignant sur les capacités générales de traitement du rythme (Activité 1 : tâches de TMS et SMS¹¹¹) et l'autre sur les capacités de traitement du langage (Activité 2 : tâche de lecture à haute voix après amorçage rythmique¹¹²) (cf. 3, p. 86).

Trois questionnements principaux ont guidé notre réflexion, que nous discutons en suivant : **l'existence d'une période de référence supramodale, l'effet de l'amorçage rythmique sur la parole et les liens entre les capacités de traitement du rythme et du langage.**

PERIODE DE REFERENCE SUPRAMODALE

Nos deux activités expérimentales nous ont permis de nous pencher sur l'existence d'une période de référence attentionnelle, individuelle, qui favoriserait le traitement d'informations temporelles dont la fréquence des pulsations est proche de cette période de référence. Pour cela, dans l'Activité 1, nous avons mesuré le Tempo Moteur Spontané (TMS) moyen des participants, représentatif de la période de référence, et la précision et la stabilité de la Synchronisation Sensori-Motrice (SMS) à différents tempi de métronome (225, 600 et 900 ms), afin d'identifier le tempo auquel la synchronisation sensori-motrice est la plus stable et/ou la plus précise. Dans l'Activité 2, nous avons relevé le Temps d'Initiation (TI) de la parole, représentatif du rythme de la parole du locuteur.

Nous avons émis l'*Hypothèse 1*, selon laquelle nous nous attendions à ce que la période de référence (TMS moyen) et le Temps d'Initiation (TI) de la parole se situent entre 400 et 800 ms. De plus, nous nous attendions à ce que les performances de précision et stabilité de la synchronisation sensori-motrice (SMS) des participants soient meilleures pour un tempo de métronome se rapprochant le plus de leur TMS moyen (période de référence) (*Hypothèse 1.1*),

¹¹¹ TMS : Tempo Moteur Spontané et SMS : Synchronisation Sensori-Motrice.

¹¹² Dans l'Activité 2, l'amorçage rythmique de la parole variait selon l'Engagement Moteur (EM) des participants, le Type d'Amorce (TA) et la Congruence Métrique (CM) entre la structure métrique sous-jacente de l'amorce et celle de la cible.

et que le temps d'initiation de la parole soit corrélé avec la mesure du TMS moyen (*Hypothèse 1.2*).

Nos résultats ont montré un TMS global (médiane des TMS moyens de l'échantillon) de 712 ms ($\sigma = 234$ ms) et un Temps d'Initiation (TI) global (moyenne des TI moyens) de 873 ms ($\sigma = 158$ ms).

Nos résultats sont cohérents avec l'existence d'une période de référence attentionnelle se situant entre 400 et 800 ms, et rejoignent ainsi les travaux de Dalla Bella et al. (2017), Drake et al. (2000) et Fraise (1974) concernant le TMS. Concernant la parole, nous avons observé un temps d'initiation moyen un peu plus lent (873 ms) que ceux généralement recensés dans la littérature (plutôt aux alentours de 600-700 ms) (Gould et al., 2016; Späth et al., 2016; Speer & Wilshire, 2014; Zhang & Zhang, 2019), et également plus lent que le temps de latence en production de mot isolé à partir de l'étape d'activation des représentations conceptuelles associées à la cible à produire, jusqu'à l'activation des représentations motrices articulatoires, de 600 ms dans le modèle de production de la parole d'Indefrey et Levelt (2004) (cf. 2.1.2.1, p. 43). Cette différence dans la durée du temps d'initiation de la parole pourrait être liée à une différence de tâche ou de matériel expérimental : le temps d'initiation de la parole est généralement observé pour de la lecture de mot isolé ou dénomination d'image (mot). Nos syntagmes expérimentaux étant linguistiquement plus complexes, les étapes de récupération lexicale et de préparation motrice pourraient prendre plus de temps. De plus, dans notre étude, le temps d'initiation de la parole est également soumis à d'autres contraintes et processus de traitement liés aux variations d'amorçage (variables indépendantes manipulées) mais aussi liés spécifiquement à la tâche de lecture.

Concernant l'*Hypothèse 1.1*, nous avons observé une plus grande difficulté à mettre en place la Synchronisation Sensori-Motrice (SMS) au tempo 225 ms, qui est le tempo le plus éloigné du TMS moyen global (712 ms), par rapport aux tempi de SMS 600 et 900 ms. En effet, parmi les séquences de SMS pour lesquelles la mise en place de la SMS n'a pas été réussie¹¹³, 92,31 % de ces séquences sont rattachées au tempo 225 ms, 7,7 % au tempo 600 ms et aucune au tempo 900 ms. Parmi les séquences de SMS présentant une asynchronie moyenne

¹¹³ *P-value* au test de Rayleigh $> 0,05$, qui illustre une répartition des frappes du sujet équivalente à une répartition de frappes aléatoires.

positive¹¹⁴, 74,34 % de ces séquences sont rattachées au tempo 225 ms, 10,26 % au tempo 600 ms et 15,38 % au tempo 900 ms. Nous avons également retrouvé une différence significative de stabilité de la SMS entre d'une part le tempo 225 ms (moins stable) et d'autre part les tempi 600 et 900 ms (plus stables) et une tendance à une précision de la SMS plus faible (asynchronie plus élevée) pour le tempo 225 ms, suivi par le tempo 600 ms et enfin 900 ms, bien qu'aucune différence de précision selon le tempo ne soit significative. De plus, les participants présentant un TMS moyen plus lent sont également ceux qui tendent à présenter une SMS la plus stable et la plus précise pour le tempo 900 ms (tempo de SMS le plus lent dans notre étude). Finalement, concernant les tempi pour lesquels les participants ont été les plus stables et/ou précis, nos résultats montrent une stabilité de la SMS la plus élevée à 600 ms pour 32,5 % des participants et à 900 ms pour 67,5 % des participants (aucun participant n'a présenté une stabilité la plus élevée pour le tempo 225 ms). Nos résultats indiquent enfin une précision de la SMS la plus élevée similaire entre les trois tempi de métronome (225 ms : 32,5 % des participants ; 600 ms : 32,5 % et 900 ms : 35 %). Nos résultats semblent ainsi aller dans le sens de l'Hypothèse 1.1 : les participants ont plus de facilité à se synchroniser à un tempo proche de leur TMS moyen, soit les tempi 600 ms et 900 ms dans notre étude. Nous retrouvons ainsi les résultats de Drake et al. (2000) et McAuley et al. (2006) concernant une facilitation de la synchronisation sensori-motrice pour des tempi se rapprochant du TMS moyen (la période de référence). De plus, tout comme Drake et al. (2000), nous relevons une préférence pour un tempo de la SMS plus lent (900 ms) que le TMS moyen (712 ms) dans le cas de deux tempi proches du TMS moyen (600 et 900 ms dans notre étude). Sur ce point, nos résultats diffèrent ainsi de ceux de Dalla Bella et al. (2017) qui ne rapportent pas de différence significative en matière de précision ou stabilité suivant le tempo de la SMS. Cette différence de résultats pourrait s'expliquer par le choix des tempi de leur étude (ISI = 450, 600 et 700 ms) qui se situent dans l'intervalle de la période de référence attentionnelle (400-800 ms).

Concernant à présent l'Hypothèse 1.2, nous n'avons pas trouvé de corrélation entre le Temps d'Initiation (TI) moyen de la parole et le TMS moyen des participants. Cependant, comme nous l'indiquions ci-dessus, nous avons mesuré le temps d'initiation de la parole en lecture de syntagmes, qui implique une planification plus longue que la production de mots isolés, les syntagmes étant linguistiquement plus complexes, ainsi que d'autres contraintes liées aux conditions d'amorçage ou encore au support écrit. Il pourrait alors être intéressant

¹¹⁴ Pour rappel, une asynchronie positive témoigne d'une réaction aux battements de métronome, à la différence d'une asynchronie négative indiquant une anticipation et donc la mise en place de la synchronisation.

d'envisager une comparaison entre le temps d'initiation de la parole en lecture de mot isolé (ou dénomination d'image) avec le TMS moyen.

A notre connaissance, notre étude est la première à questionner un lien potentiel entre le TMS moyen et le Temps d'Initiation (TI) de la parole. L'absence de corrélation dans nos résultats entre le TMS et le TI suggère que différents mécanismes sous-tendent l'établissement de la période de référence pour les gestes manuels et les gestes d'initiation de la parole pour la lecture de syntagmes. Cette mise en relation du TMS et du TI serait cependant à explorer dans une tâche de parole sans paradigme d'amorçage (plus spontanée), ou avec du matériel linguistiquement moins complexe, comme par exemple une tâche de lecture de mot isolé ou de dénomination d'image. Nous pensons également qu'il pourrait être intéressant de comparer le TMS moyen et l'intervalle inter-accentuel moyen de la parole. En effet, comme nous l'avons présenté dans la partie 1.2.1.3 (p. 19), l'intervalle inter-accentuel moyen se situe aux alentours de 600 ms dans toutes les langues, et pourrait ainsi refléter la période de référence attentionnelle. A la différence du temps d'initiation qui est une mesure isolée de réaction à la présentation d'un stimulus (par exemple, phrase à lire, image à dénommer), l'intervalle inter-accentuel reflète la pulsation sous-jacente dans le flux de parole, ce qui en ferait un meilleur candidat pour la comparaison d'une période de référence supramodale avec le TMS moyen. Cette question d'une relation entre l'intervalle inter-accentuel et le TMS moyen pourra être investiguée à partir des enregistrements de syntagmes recueillis lors de l'Activité 2, que nous avons segmentés partiellement pour obtenir la mesure du temps d'initiation de la parole. La poursuite de la segmentation et de l'annotation des enregistrements nous permettra d'obtenir les mesures des intervalles inter-accentuels.

Nos résultats montrent également une certaine **incidence du profil socio-démographique** sur le tempo moteur spontané moyen, les tempi les plus stables et/ou précis pour la synchronisation sensori-motrice et le temps d'initiation de la parole. Nous précisons cependant que le profil socio-démographique n'a pas été contrôlé spécifiquement pour cette étude, les observations que nous en tirons sont donc à prendre avec précaution.

Nos résultats montrent un ralentissement significatif du TMS moyen avec l'avancée de l'âge (tout comme Dalla Bella et al., 2017; Drake et al., 2000; Fraise, 1974; McAuley et al., 2006) ainsi qu'une tendance à une accélération du TMS moyen avec l'augmentation des années

de pratique de la danse. Nous avons également relevé d'autres tendances, comme un TMS moyen plus rapides chez les femmes, les musiciens (et notamment les percussionnistes), les danseurs et les joueurs des jeux vidéo¹¹⁵.

Concernant la synchronisation sensori-motrice, nos résultats vont dans le sens de ceux de Drake et al. (2000) en suggérant que les musiciens tendent à présenter une plus grande étendue dans la répartition des tempi les plus stables et précis possibles pour la SMS, ainsi que dans le sens des travaux de Drake et al. (2000) et Thompson et al. (2015) qui suggèrent que les musiciens sont plus stables et plus précis que les non-musiciens et que ces mesures s'améliorent avec l'augmentation des années de pratique de la musique.

Le Temps d'Initiation (TI) moyen de la parole est aussi sujet aux variations individuelles : tout comme le TMS moyen, le TI tend à ralentir avec l'âge mais également avec l'augmentation du nombre d'années de pratique de la musique. De plus, le temps d'initiation moyen tend à être plus rapide chez les participants joueurs de jeux vidéo.

Ces résultats à propos de l'effet de l'âge ou de la pratique des jeux vidéo, bien que provenant de VI non contrôlées spécifiquement, invitent cependant à prendre en considération certaines caractéristiques socio-démographiques, notamment l'âge ou le niveau d'expertise rythmique (nombre d'années de pratique de la danse ou de la musique). De plus, ces résultats nous confortent dans nos choix de critères de sélection des participants : nous avons sélectionné uniquement des participants adultes et non professionnels de la danse ou de la musique. Une tranche d'âge plus étendue ou l'intégration de profils experts rythmiques plus variés pourraient permettre de caractériser des différences individuelles plus importantes, concernant le TMS moyen, les tempi les plus stables et/ou précis et le temps d'initiation de la parole, ainsi que leur mise en comparaison.

AMORÇAGE RYTHMIQUE DE LA PAROLE

L'objectif principal de notre étude était d'identifier les modalités rythmiques les plus efficaces pour faciliter l'initiation de la parole, afin de proposer des pistes pour la rééducation de troubles de la parole ou du langage. Nous avons ainsi observé le mécanisme de synchronisation et d'*entraînement* de la parole sur un stimulus rythmique extérieur dans la

¹¹⁵ Pour rappel, dans cette étude, nous employons les termes « musicien », « danseur » ou « joueur de jeux vidéo » dès lors qu'un participant a indiqué pratiquer ou avoir pratiqué une activité musicale, de danse ou de jeux vidéo, respectivement.

deuxième activité expérimentale *via* le paradigme de l'amorçage, que nous avons investigué en faisant varier différents paramètres : l'Engagement Moteur (EM) du participant sur l'amorce, le Type d'Amorce (TA) et la Congruence Métrique (CM) entre la structure métrique du stimulus d'amorce et celle du syntagme cible. L'effet des variations d'amorçage rythmique a été observé sur le nombre d'erreurs de lecture produites et le temps d'initiation de la parole, ces deux variables renseignant sur le caractère plus ou moins facilitant des différents amorçages sur la production de parole : un amorçage associé à un nombre d'erreurs le plus bas et un TI le plus court étant considéré, par hypothèse, comme le plus facilitant.

Ainsi, nous nous attendions à ce que la condition d'amorçage avec engagement moteur (+EM) soit plus facilitante pour la production de parole que la condition d'amorçage sans engagement moteur (-EM). En effet, l'existence d'un partage anatomo-fonctionnel entre les aires de contrôle des gestes manuels et des gestes de production de la parole se traduirait par une facilitation de l'initiation de la parole suite à la production d'un geste manuel (*Hypothèse 2*).

Ensuite, concernant le Type d'Amorce, nous nous attendions à ce que la condition d'amorçage de type vocal langagier (L) soit plus facilitante que la condition d'amorce vocale non langagière (NL), elle-même plus facilitante que la condition d'amorce musicale (M). En effet, la perception d'amorces vocales (langagières et non langagières) convoque le système moteur spécifique à la production de la parole. Ces amorces, plus incarnées, seraient ainsi facilitantes par rapport à une amorce musicale. En effet, selon une perspective incarnée, notre cognition repose sur nos habiletés et expériences perceptives et motrices (Iverson & Thelen, 1999), supposant un couplage sensori-moteur fort entre perception et production de parole/vocale. Deuxièmement, nous nous attendions à ce que le traitement de l'amorçage langagier soit automatique et stimule le traitement lexical, facilitant à la fois le traitement perceptif de l'amorce et la production lexicale cible subséquente, par rapport à l'amorçage vocal non langagier (*Hypothèse 3*).

Par ailleurs, nous nous attendions à ce que la condition d'amorçage métriquement congruent (C) soit plus facilitante que la condition d'amorçage métriquement incongruent (IC). En effet, en vertu de la Théorie Dynamique de l'Attention (DAT), les rythmes attentionnels étant synchronisés sur la structure métrique de l'amorce, cette structure devrait persister (être entraînée) et servir de patron métrique pour la production du syntagme cible subséquent. Ainsi, si les structures de l'amorce et de la cible sont métriquement congruentes, la production de

parole est facilitée, tandis qu'en situation d'amorçage incongruent, le couplage oscillatoire doit être défait pour instaurer un nouveau patron métrique sous-jacent, affectant la production du syntagme cible subséquent (*Hypothèse 4*).

Finalement, nous avons émis l'hypothèse que la congruence métrique de l'amorçage interagit avec l'engagement moteur du participant. En effet, la production d'un geste moteur co-occurent avec les pulsations d'un stimulus rythmique facilite l'extraction de la pulsation et renforce donc la perception de la structure métrique sous-jacente de l'amorce. Ainsi, en situation d'amorçage métriquement congruent, l'ajout de l'engagement moteur devrait faciliter l'instauration du patron métrique (congruent) de la cible subséquente et donc faciliter fortement la production de parole. Au contraire, en situation d'amorçage métriquement incongruent, l'ajout de l'engagement moteur, en renforçant les attentes métriques devrait affecter fortement la production de parole subséquente (*Hypothèse 5*).

Nous discutons ci-dessous ces hypothèses au regard de nos résultats.

EFFET DE L'ENGAGEMENT MOTEUR DU PARTICIPANT LORS DE L'AMORÇAGE

Dans notre étude, l'amorçage rythmique avec Engagement Moteur (+EM) se traduit par une réduction du Temps d'Initiation (TI) de la parole par rapport à un amorçage sans Engagement Moteur (-EM), indiquant un effet facilitateur pour la production de parole suite à une stimulation motrice manuelle. Ce résultat est cohérent avec l'existence d'un partage anatomo-fonctionnel entre la motricité manuelle et les gestes de production de la parole (Dick et al., 2019; Gentilucci & Corballis, 2006; Iverson & Thelen, 1999), et rejoint ainsi les résultats de Falk et Dalla Bella (2016) qui montrent un avantage de l'amorçage rythmique audio-moteur pour la *perception* de la parole par rapport à un amorçage rythmique auditif seul. Notre étude se différencie cependant en ce qu'elle interroge le versant de la *production* de la parole, ce résultat permettant de renforcer l'argument d'un couplage fort sensorimoteur entre le versant de la perception et de la production de la parole. A propos des différences individuelles, nos résultats indiquent cependant que cet effet de l'engagement moteur pourrait être modulé par la pratique musicale, seuls les participants musiciens présentant un effet facilitateur de l'amorçage rythmique auditif avec engagement moteur (+EM) par rapport à l'amorçage rythmique sans engagement moteur (-EM).

De plus, le fait que l'effet significatif de l'engagement moteur sur le temps d'initiation de la parole soit retrouvé quel que soit le Type d'Amorce (TA) renforce l'hypothèse de mécanismes partagés pour le traitement de stimuli musicaux (M) et vocaux (langagiers – L et non langagiers – NL). En ce sens, nos résultats sont cohérents avec (1) l'hypothèse OPERA (Patel, 2011), selon laquelle le traitement de la musique et le traitement de la parole reposent sur un mécanisme partagé pour le traitement temporel fin (régime spectral et régime temporel) par le système auditif, ainsi qu'avec (2) les propositions de Drake (1998) et Di Cristo (2003) qui considèrent la parole comme un signal rythmique, ceci impliquant que sa perception se fasse de la même manière que le traitement des séquences auditives complexes telles que la musique (segmentation du signal en groupes, extraction de la pulsation, et intégration de ces deux premiers processus dans des structures hiérarchiques).

L'analyse du nombre d'erreurs de lecture nous montre cependant que le raccourcissement des temps d'initiation associé à un amorçage rythmique auditif avec engagement moteur (+EM) pour les amorces de type langagier (L) et non langagier (NL) peut augmenter la probabilité de produire une erreur de lecture par rapport à un amorçage sans engagement moteur (-EM), bien que cette probabilité reste faible. En effet, le temps d'initiation de la parole étant raccourci, le temps alloué au processus de contrôle anticipatoire de production de la parole (l'étape de programmation motrice, que l'on retrouve dans le modèle d'Indefrey et Levelt, 2004 ou le modèle DIVA de Tourville et Guenther, 2011) est réduit, ce qui pourrait affecter la qualité (précision) des processus de monitoring (Postma, 2000) et se traduire par une augmentation du nombre d'erreurs de lecture. En résumé, même si ce compromis entre vitesse et précision semble ressortir ici, quoique de façon marginale, la production de la parole et le temps de latence associé dépendent également de processus de contrôle exécutif.

Concernant les interactions entre motricité et attention/perception, il pourrait être intéressant d'aller plus loin pour distinguer (1) la part de la motricité liée au partage anatomo-fonctionnel entre la motricité manuelle et les gestes de production de la parole et (2) la part de la motricité liée à une amélioration de l'attention et donc de la perception rythmique. Les travaux de Morillon et al. (2014) à propos de l'influence de la motricité sur l'attention pointent sur la nécessité d'une co-occurrence entre les mouvements rythmiques manuels et le stimulus rythmique pour améliorer la perception (stimulus rythmique musical), tandis que les travaux de Falk et Dalla Bella (2016) indiquent que les mouvements manuels, qu'ils soient co-occurents (métriquement congruents) ou non avec la structure de l'amorce rythmique (stimulus rythmique langagier dans leur étude), sont avantageux pour la détection d'un changement lexical

subséquent par rapport à une condition d'amorçage sans engagement moteur ou une condition sans amorçage rythmique, cet avantage étant renforcé lorsque les mouvements manuels et l'amorce rythmique langagière sont alignés. L'avantage de la condition d'un alignement incongruent entre les mouvements manuels et l'amorce rythmique langagière sur une condition sans amorçage n'est cependant pas retrouvé dans une seconde étude de Falk et al. (2017). Nous pourrions ainsi envisager dans une future étude d'intégrer une condition supplémentaire concernant la congruence métrique entre les mouvements manuels et l'amorce rythmique, cet effet n'ayant à notre connaissance pas été testé sur le versant de la production du langage : dans notre étude, lors de la condition avec Engagement Moteur (+EM), les mouvements manuels étaient alignés sur les battements de l'amorce rythmique.

A propos des différences individuelles, nos résultats indiquent cependant que cet effet de l'engagement moteur pourrait être modulé par la pratique musicale, puisque seuls les participants ayant indiqué pratiquer ou avoir pratiqué une activité musicale présentent un effet facilitateur de l'amorçage rythmique auditif avec engagement moteur par rapport à un amorçage sans engagement moteur.

EFFET DU TYPE D'AMORCE

Nos résultats indiquent qu'un amorçage de type vocal non langagier (NL) provoque une réduction significative du temps d'initiation de la parole par rapport à un amorçage vocal langagier (L). L'effet de l'amorçage musical (M) sur le temps d'initiation de la parole ne se distingue pas significativement des effets des amorces de type langagier et non langagier. Cependant, nos résultats montrent une tendance à une réduction du temps d'initiation de la parole suite à un amorçage non langagier par rapport à un amorçage musical, et une tendance à une réduction du temps d'initiation de la parole suite à un amorçage musical par rapport à un amorçage vocal langagier. Nous ne retrouvons ainsi pas tout à fait les résultats concernant l'effet du type d'amorce sur le temps d'initiation de la parole que nous avons observé lors d'une analyse précédente sur les mêmes données (Gindre et al., 2023), qui indiquaient une différence significative entre d'une part le type d'amorçage non langagier et d'autre part les amorçages langagier et musical. Cette différence s'explique par le fait que nous avons été plus conservatrice dans notre approche d'analyse statistique, en prenant la décision d'exclure de l'analyse les données d'un participant dont le profil se différenciait du reste de l'échantillon (cf.

5.1, p. 109) et en appliquant une approche de modélisation linéaire plus stricte basée sur la comparaison et la sélection des effets aléatoires et des effets fixes.

Au lieu d'indiquer une facilitation de l'amorçage langagier par rapport aux amorçages non langagier et musical comme nous l'avons envisagé avec l'*Hypothèse 3* (effet envisagé : temps d'initiation de la parole langagier plus rapide (court) < non langagier < musical), nos résultats semblent plutôt indiquer que les amorces de type non langagier et musical sont plus facilitantes pour la production de parole qu'une amorce de type langagier. Nous nous attendions à ce que l'effet de lexicalité favorise le traitement des mots par rapport aux non-mots (dans les amorces) et facilite la production lexicale subséquente. Cependant, nos résultats montrent un avantage des amorces non langagières (avantage significatif) et musicale (tendance) par rapport à l'amorce langagière, qui pourrait s'expliquer par le fait que les amorces non langagières et musicale sont neutres en termes de ressources dédiées au processus de traitement des informations sémantiques et syntaxiques, en perception. Ainsi, elles ne créent pas de contexte préalable à la lecture des syntagmes cibles. Au contraire, les amorces de type langagier installent un contexte sémantique et syntaxique qui viendrait interférer avec la production du syntagme cible subséquent, et qui impliquent des traitements d'informations linguistiques demandant un temps de traitement supplémentaire, par rapport aux autres types d'amorces (non langagier et musical).

La tendance d'un effet facilitateur de l'amorce non langagière par rapport à l'amorce musicale, bien que ces deux types d'amorces induisent autant un contexte neutre l'une que l'autre par rapport à l'amorce langagière, suggère l'activation des aires motrices de production de la parole lors de l'écoute des amorces non langagières. Cette observation est cohérente avec les approches sensori-motrices de perception de la parole, tel que le modèle neuropsycholinguistique de Hickok et Poeppel (2004, 2007). L'initiation motrice de la parole serait ainsi facilitée suite à cette stimulation par rapport à l'amorce musicale qui ne provoque pas cette activation. Ces résultats sont intéressants à envisager dans un contexte de remédiation rythmique de troubles de la parole et notamment des difficultés d'initiation de la parole, par exemple pour le bégaiement ou l'aphasie non fluente. Il pourrait ainsi être pertinent d'adapter cette méthodologie auprès de ces populations de patients afin de tester l'effet de l'amorçage rythmique non langagier par rapport à l'amorçage langagier ou musical sur la facilitation de l'initiation de la parole.

L'analyse du nombre d'erreurs de lecture n'a pas montré d'effet différentiel significatif selon le type d'amorce. Cela pourrait s'expliquer notamment par le nombre faible d'erreurs (4,48 % des syntagmes enregistrés présentent des erreurs de lecture).

EFFET DE LA CONGRUENCE METRIQUE

La manipulation de la congruence métrique (CM) nous a permis d'observer les mécanismes de synchronisation et *entrainement* de la parole sur un stimulus externe métrique, soit sur de multiples périodicités (associées chacune à un niveau prosodique) liées entre elles par des relations métriques hiérarchiques. La synchronisation des rythmes attentionnels, et donc de la parole rythmique produite, sur les périodicités d'un rythme externe complexe (métrique) rejoint la Théorie dynamique de l'attention (DAT), ici appliquée à la parole. Dans notre étude, nous avons considéré trois niveaux hiérarchiques de périodicités (du plus bas au plus haut) : le niveau syllabique, le niveau du syntagme accentuel (le plus souvent, dans notre corpus, identique au niveau du mot prosodique), associé à l'accent final (AF) de syntagme, et le niveau du syntagme intonatif, associé à un AF d'unité intonative. Nous avons manipulé la congruence métrique entre la structure du stimulus d'amorce et celle du syntagme cible, en associant une amorce et une cible présentant ou non la même structure métrique, soit appartenant ou non au même Groupe Métrique (GM).

La différence de congruence métrique entre la structure métrique du stimulus d'amorce et celle du syntagme cible n'a pas produit d'effet différentiel significatif tel que prédit par l'*Hypothèse 4*, que ce soit sur la précision (nombre d'erreurs de lecture) ou sur le temps d'initiation de la parole. Nos résultats ne permettent donc pas de répliquer les observations de Cason, Hidalgo, et al. (2015), Contreras Roa et al. (2022), Gould et al. (2016) ou Zhang et Zhang (2019) à propos de l'avantage d'un amorçage métriquement congruent sur la production de parole (précision phonologique ou temps d'initiation de la parole) chez des enfants sourds et des adultes neurotypiques et parkinsoniens.

Cependant, comme indiqué par Zhang et Zhang (2019) pour le versant de la production de la parole et Cason, Astésano, et al. (2015) pour le versant de la perception, la structure rythmique de l'amorce et de la cible semble avoir une importance dans la sensibilité à un effet de la congruence métrique. En effet, Zhang et Zhang (2019) n'ont trouvé un effet significatif de la congruence métrique que pour un groupe métrique de cible sur deux, et Cason, Astésano, et al. (2015) n'ont trouvé un effet de la congruence métrique que pour deux groupes métriques

d'amorces sur quatre. Zhang et Zhang (2019) suggèrent ainsi que la sensibilité à la congruence métrique pourrait être dépendante de la fréquence du patron métrique abstrait dans la langue étudiée. Nous avons alors complété notre modèle inférentiel initial de sorte à observer les interactions possibles entre la congruence métrique et le groupe métrique de l'amorce. Nos résultats semblent ainsi indiquer un effet facilitant de l'amorçage métriquement congruent, mais uniquement pour le Groupe Métrique 2 (x x X x x X), tandis qu'un effet facilitant de l'amorçage incongruent est retrouvé pour le Groupe Métrique 4 (x x x X x x X), et aucun effet n'est retrouvé pour les amorces des Groupes Métriques 1 (x x x X x X) et 3 (x x X x x x X). L'observation d'un effet contradictoire de la congruence métrique entre les amorces des Groupes Métrique 2 et 4 dans notre étude pourrait alors plutôt indiquer un effet du Groupe Métrique de la cible plutôt qu'un effet du Groupe Métrique de l'amorce, avec une cible du Groupe Métrique 2 qui serait plus avantageuse qu'une cible du Groupe Métrique 4 sur l'initiation de la parole. De plus, le Groupe Métrique 2 présente une structure répétitive de trois éléments (x x X x x X). Selon Fraise (1974), nous avons tendance à grouper naturellement les éléments par deux ou trois et plus rarement par quatre (cf. 1.2.1.2, p. 17), ce qui pourrait indiquer un ancrage plus fort des schémas métriques comportant trois éléments. Cela se traduirait à la fois par une plus grande inertie métrique lorsque la cible est du Groupe Métrique 2 (peu importe le Groupe Métrique de l'amorce), et à la fois par des attentes métriques plus fortes relatives à cette structure et une plus grande sensibilité à l'effet de congruence lorsque l'amorce est du Groupe Métrique 2. Une poursuite de ce questionnement pourrait être d'envisager une comparaison de l'effet de la Congruence Métrique de différents Groupes Métriques (amorces et cibles) constitués de sous-structures répétitives (par exemple les Groupes Métriques suivants : x X x X ; x x X x x X ; x x x X x x x X), sur la production de parole.

Nous ne retrouvons donc que partiellement les observations de Cason, Astésano, et al. (2015), qui avaient montré un effet facilitant de l'amorçage congruent pour les amorces des Groupes Métriques 2 et 3, sur le versant de la perception de la parole. Cependant, il faut noter qu'afin de nous concentrer sur nos trois effets d'intérêt principaux (engagement moteur, type d'amorce et congruence métrique), nous n'avons pas envisagé d'observer l'effet du groupe métrique de la cible dans cette étude et n'avons de fait pas considéré autant de niveaux de granularité de congruence dans notre protocole que dans l'étude de Cason, Astésano, et al. (2015). En effet, nous n'avons considéré qu'une seule situation d'incongruence, soit lorsqu'une amorce et une cible présentent à la fois une différence d'emplacement du premier accent final et un nombre d'éléments différents, soit les combinaisons incongruentes suivantes : Groupe

Métrique 1 avec Groupe Métrique 3 (et inversement) et Groupe Métrique 2 avec Groupe Métrique 4 (et inversement). Une poursuite de ce questionnement pourrait être de répliquer la méthodologie de Cason, Astésano, et al. (2015) avec davantage de niveaux d'incongruence (incongruence uniquement sur l'emplacement du premier accent final de syntagme ; incongruence uniquement sur le nombre de syllabes ; incongruence complète), appliquée cette fois sur le versant de la production de la parole.

Une perspective de cette étude concerne justement la poursuite de la segmentation et de l'annotation des enregistrements des syntagmes cibles de l'Activité 2 (uniquement segmentés pour recueillir la mesure du temps d'initiation pour l'instant), en annotant la structure métrique sous-jacente réalisée. Notamment, l'analyse des productions cibles suite à un amorçage métriquement incongruent pourra nous renseigner sur le poids de l'ancrage métrique et temporel de la structure de l'amorce. En effet, dans notre étude, l'amorçage incongruent se traduit par une différence d'emplacement du premier accent final de syntagme ainsi que par une différence du nombre d'éléments¹¹⁶ entre la structure de l'amorce et de la cible, soit également une différence de durée. Ainsi, dans le cas d'un amorçage incongruent, la structure métrique cible produite pourrait présenter un déplacement du premier accent final (ajustement métrique) et/ou une adaptation des durées syllabiques traduisant une différence du nombre de syllabes entre la structure de l'amorce et la structure de la cible (ajustement temporel). Nous faisons l'hypothèse que, du fait de l'importance de la structure métrique pour l'accès au sens, l'ancrage métrique prime sur l'ancrage temporel. Ainsi, si nous prenons l'exemple d'un amorçage métriquement incongruent impliquant une amorce du Groupe Métrique 2 (x x X x x X) et une cible du Groupe Métrique 4 (x x x X x x X), nous nous attendons à voir dans la réalisation métrique cible un déplacement du premier accent final en troisième position (et non plus en quatrième), plutôt qu'un rééquilibrage temporel des durées des syllabes de la cible pour calquer la structure temporelle de l'amorce.

EFFETS DE L'INTERACTION ENTRE LA CONGRUENCE METRIQUE ET L'ENGAGEMENT MOTEUR

Nous n'avons pas trouvé d'interaction entre la congruence métrique et l'engagement moteur des participants, conformément à ce nous proposons dans l'*Hypothèse 5*. Nos résultats ne permettent ainsi pas de répliquer (sur le versant de la production) les observations de Cason,

¹¹⁶ Éléments = sons de percussions pour les amorces musicales, ou bien syllabes pour les amorces langagières et non langagières ainsi que pour les syntagmes cibles.

Astésano, et al. (2015), qui indiquent avoir trouvé un effet de la congruence métrique entre l'amorce et la cible sur la perception de la parole (détection de phonème dans des syntagmes) renforcé (effet significatif de la congruence métrique) pour un groupe de participants ayant suivi un court exercice d'engagement moteur (copie vocale de l'amorce rythmique musicale), par rapport à un groupe de participants ne l'ayant pas suivi (effet non significatif de la congruence métrique).

En résumé pour l'Activité 2, la survenue d'effets différentiels dans le nombre d'erreurs et le temps d'initiation de la parole suite aux variations d'amorçage (engagement moteur, type d'amorce et congruence métrique) indique :

- une facilitation de l'amorçage rythmique auditif avec engagement moteur du participant (+EM) par rapport à un amorçage rythmique auditif seul (effet significatif), ce qui est cohérent avec l'existence d'un partage anatomo-fonctionnel entre les systèmes de contrôle des mouvements de la main et des mouvements impliqués dans la production de la parole. Ainsi, la production d'un geste manuel stimulerait les aires motrices de production de la parole, facilitant la production de parole subséquente. De plus, l'effet de l'engagement moteur est significatif quel que soit le type d'amorce (langagier, non langagier ou musical), ce qui est cohérent avec l'existence de mécanismes supramodaux partagés entre le traitement de stimuli rythmiques musicaux, vocaux non-langagiers et langagiers : les amorces langagières, non langagières et musicales sont toutes trois des séquences auditives complexes, dont le décodage repose sur des mécanismes cognitifs partagés ;
- un effet facilitant de l'amorçage rythmique non langagier (NL) par rapport à un amorçage musical (effet non significatif, dans notre modèle statistique conservateur actuel), ce qui est cohérent avec les modèles de production de la parole impliquant les mécanismes de perception de la parole lors de la préparation motrice ;
- un effet facilitant de l'amorçage rythmique non langagier par rapport à un amorçage langagier (effet significatif), ce qui indique une facilitation d'un amorçage neutre d'un point de vue sémantique et syntaxique par rapport à un amorçage langagier qui instaure un contexte préalable à la lecture des cibles et vient interférer avec la production de parole subséquente ;

- une facilitation de l'amorçage métriquement congruent (C) pour le Groupe Métrique 2, ce qui rejoint les résultats de Cason, Astésano, et al. (2015) qui avaient également trouvé un effet de la congruence métrique pour une amorce de ce groupe métrique. D'autre part, ce résultat pourrait s'expliquer par la structure répétitive de trois éléments (x x X x x X) de ce groupe métrique, qui favorise un ancrage métrique plus fort que les structures non répétitives ou comportant plus d'éléments, le groupement s'établissant plus naturellement sur deux ou trois éléments (Frasse, 1974).

Il pourrait être intéressant d'adapter cette méthodologie auprès de populations présentant des troubles de la fluence verbale et notamment de l'initiation motrice de la parole, soit des pathologies telles que le bégaiement, l'aphasie, l'apraxie ou encore la dysarthrie parkinsonienne. En effet, des protocoles de rééducation intègrent déjà la stimulation rythmique auditive ou audio-motrice comme support pour la production de parole (cf. 2.3.2.2, p. 65) et pourraient de fait bénéficier de recherches sur l'effet différentiel de variations d'amorçage afin de proposer un protocole plus adapté, en particulier concernant l'Engagement Moteur du participant et le type d'amorce Non Langagier.

Enfin, il pourrait être intéressant d'observer l'effet de l'amorçage rythmique de la production de parole en fonction de la tâche. En effet, dans notre étude, nous avons investigué l'effet de l'amorçage rythmique sur la production de la parole dans une tâche de lecture à haute voix. Le support écrit est particulier en ce qu'il stimule les aires cérébrales de production de la parole, et ce même en lecture silencieuse (Indefrey & Levelt, 2004; Sun & Peperkamp, 2016). Par exemple, une tâche de lecture de mots isolés à haute voix pourrait ainsi être facilitante pour initier la production de la parole par rapport à une tâche de production orale (mot) à partir d'une image.

LIENS ENTRE LES CAPACITES DE TRAITEMENT DU RYTHME ET DU LANGAGE

S'agissant du partage des mécanismes de traitement du rythme et de traitement du langage, nous avons mis en regard les observations de l'Activité 1 et de l'Activité 2. Nous nous attendions à ce que le nombre d'erreurs et le Temps d'Initiation (TI) de la parole (Activité 2) soient moins élevés chez les participants présentant une stabilité plus élevée du Tempo Moteur Spontané (TMS) et de la Synchronisation Sensori-Motrice (SMS) et une précision plus élevée de la SMS (cf. Hypothèse 6).

Concernant le temps d'initiation de la parole, celui-ci est corrélé positivement à la stabilité de la SMS au tempo 600 ms : les participants les plus stables pour la SMS au tempo 600 ms présentent un temps d'initiation de la parole plus lent. Nous observons de plus une tendance à un ralentissement du temps d'initiation avec l'augmentation de la stabilité et de la précision de la SMS, tous tempi confondus. Le temps d'initiation de la parole pourrait ainsi être un indicateur des capacités de traitement du rythme, cependant, dans un sens contraire à nos attentes : le temps d'initiation de la parole ralentit avec l'augmentation de la stabilité du TMS et de la SMS et de la précision de la SMS (nous discutons ce point ci-après).

Concernant la précision de la production de parole, nous avons observé que le nombre d'erreurs de lecture produites tend à diminuer avec l'augmentation de la stabilité et de la précision moyennes (tous tempi confondus) de la SMS, et ce notamment pour la SMS au tempo 225 ms : les participants les plus stables et précis dans la tâche de SMS sont ceux qui présentent le moins d'erreurs de lecture.

Nos résultats semblent ainsi indiquer que les capacités de traitement rythmique sont reflétées à la fois par les mesures comportementales telles que le temps d'initiation de la parole et le nombre d'erreurs, les participants les plus stables et précis lors de l'Activité 1 étant ceux présentant le temps d'initiation de la parole le plus lent et ceux qui produisent le moins d'erreurs. Nous pourrions interpréter cette différence concernant le temps d'initiation par rapport à nos attentes comme le reflet d'un contrôle attentionnel accru, qui retentit ainsi à la fois sur un temps d'initiation plus lent et sur un nombre d'erreurs plus faible (Levelt et al., 1999). Ce questionnement pourrait être investigué dans une tâche linguistique faisant varier le degré de contrôle attentionnel (focalisation attentionnelle, attention soutenue ou attention divisée), afin d'en mesurer l'impact différentiel sur le temps d'initiation de la parole et le nombre d'erreurs et de les mettre en regard à nouveau des capacités de traitement du rythme.

De plus, nous avons relevé un effet significatif du nombre d'années de pratique de la musique sur la stabilité et la précision de la SMS au tempo 225 ms, indiquant que les participants présentant un niveau d'expertise musicale plus élevé ont de meilleures performances de stabilité et précision de la SMS au tempo 225 ms. La pratique musicale pourrait ainsi avoir une incidence sur les capacités de traitement du rythme. Nous rappelons cependant que le profil socio-démographique n'a pas été contrôlé spécifiquement pour cette étude, les observations que nous en tirons sont donc à prendre avec précaution.

POUR CONCLURE, notre étude nous a permis de nous pencher sur l'existence d'une période de référence supramodale, l'effet de l'amorçage rythmique sur la parole et les liens entre les capacités de traitement du rythme et du langage.

Pour investiguer l'existence d'une période de référence supramodale, nous avons recueilli et mis en regard différentes mesures rythmiques : le Tempo Moteur Spontané (TMS) moyen et le tempo de Synchronisation Sensori-Motrice (SMS) le plus stable et/ou le plus précis dans l'Activité 1, et le Temps d'Initiation (TI) de la parole dans l'Activité 2. Le TMS moyen global de notre échantillon est de 712 ms et le TI moyen global est de 873 ms. Nous avons pu observer un lien entre le TMS moyen et les tempi les plus stables et précis de la synchronisation sensori-motrice, mais pas entre le TMS moyen et le temps d'initiation de la parole. Ces résultats confortent l'existence d'une période de référence supramodale aux alentours de 400-800 ms, bien que différents mécanismes semblent sous-tendre l'établissement de la période de référence pour les gestes manuels et les gestes d'initiation de la parole pour la lecture de syntagmes.

Concernant l'effet de l'amorçage rythmique sur la facilitation de la production de parole, nous avons trouvé un effet facilitant significatif d'un amorçage rythmique auditif avec Engagement Moteur (+EM) du participant par rapport à un amorçage rythmique sans Engagement Moteur (-EM), ainsi qu'un effet facilitant significatif d'un amorçage rythmique vocal Non Langagier (NL) par rapport à un amorçage rythmique vocal Langagier (L), et une tendance à un effet facilitant d'un amorçage rythmique non langagier par rapport à un amorçage rythmique musical (M), ce dernier tendant à être plus facilitant que l'amorçage rythmique langagier (soit la tendance suivante : temps d'initiation NL plus rapide (court) < M < L). Nous n'avons pas trouvé d'effet principal de la Congruence Métrique (CM) de l'amorçage, ni en interaction avec l'engagement moteur du participant. Nous avons cependant trouvé une interaction entre la condition de Congruence Métrique et le Groupe Métrique de l'amorce. Ces résultats encouragent à poursuivre les recherches sur l'effet de l'amorçage rythmique avec engagement moteur et/ou avec des amorces de type non langagier auprès de populations présentant des troubles de la production de la parole et notamment une difficulté de l'initiation de la parole.

Enfin, à propos des liens entre les capacités de traitement du rythme et du langage, nous avons trouvé que les participants les plus stables dans la tâche de SMS au tempo 600 ms sont les participants dont le temps d'initiation de la parole est le plus lent (corrélation significative), ainsi qu'une tendance générale indiquant que les participants les plus stables et les plus précis

dans la tâche de SMS (tous tempi confondus) sont également les participants dont le temps d'initiation est le plus lent et ceux qui présentent le moins d'erreurs de lecture. D'un point de vue du contrôle attentionnel, ces résultats semblent indiquer qu'un contrôle accru est reflété par un temps d'initiation de la parole plus lent, ce qui permet de produire moins d'erreurs de lecture. Ces résultats invitent à poursuivre les recherches à la croisée des domaines sur les mécanismes de traitement partagés entre rythme et langage, dans le but de mieux intégrer les mesures des capacités de traitement du rythme dans l'évaluation des capacités de traitement du langage (et inversement), notamment à travers l'étude des temps de latences, des phénomènes de (dis)fluences au sens large, et des paramètres métriques en perception et production de la parole typique et atypique.

Enfin, nous avons observé une certaine incidence du profil socio-démographique des participants dans la réalisation des différentes tâches expérimentales, notamment en ce qui concerne l'âge ou la pratique de la musique. Bien que ces informations aient été recueillies à titre indicatif sans avoir été contrôlées et donc équilibrées dans la présente étude, nos observations encouragent à les prendre en considération. La prise en compte des caractéristiques socio-démographiques, notamment l'âge et l'expertise rythmique (pratique de la musique et de la danse), nous semble ainsi importante à intégrer et contrôler pour ce type de méthodologie, que ce soit au moment de la constitution de l'échantillon de participants ou dans les analyses statistiques. Par exemple, nos résultats ont montré que l'effet de l'engagement moteur sur le temps d'initiation de la parole est modulé par la pratique musicale, seuls les participants ayant indiqué pratiquer ou avoir pratiqué une activité musicale étant sensibles à cet effet d'amorçage dans notre étude. Il pourrait être intéressant de répliquer cette méthodologie en testant l'effet de l'engagement moteur auprès d'une population dont la pratique musicale est davantage contrôlée, par exemple avec deux groupes, l'un composé de participants musiciens professionnels et l'autre de participants non-musiciens. Ces résultats seraient pertinents pour pouvoir ensuite discuter de l'utilisation de l'amorçage rythmique avec engagement moteur dans la remédiation de pathologies de la parole ou du langage en tenant compte de l'expertise musicale des patients.

BIBLIOGRAPHIE

- Aguilera, M., El Yagoubi, R., Espesser, R., & Astésano, C. (2014). Event-Related Potential investigation of Initial Accent processing in French. *Proceedings of the International Conference on Speech Prosody, May*. <https://doi.org/10.21437/speechprosody.2014-63>
- Aichert, I., Lehner, K., Falk, S., Späth, M., Franke, M., & Ziegler, W. (2021). In time with the beat: Entrainment in patients with phonological impairment, apraxia of speech, and parkinson's disease. *Brain Sciences, 11*(11). <https://doi.org/10.3390/brainsci11111524>
- Albert, A., & Grice, M. (n.d.). Rhythm is a timescale. In L. Meyer & A. Strauss (Eds.), *Rhythms of Speech and Language: Culture, Cognition, and the Brain*. Cambridge University Press. <https://doi.org/10.31234/osf.io/h3g7t>
- Albert, M. L., Sparks, R. W., & Helm-Estabrooks, N. A. (1973). Melodic Intonation Therapy for Aphasia. *Archives of Neurology, 29*(2), 130–131.
- Arvaniti, A. (2009). Rhythm, timing and the timing of rhythm. *Phonetica, 66*(1–2), 46–63. <https://doi.org/10.1159/000208930>
- Astésano, C. (2001). *Rythme et Accentuation en Français : Invariance et Variabilité Stylistique*. L'Harmattan.
- Astésano, C. (2014). *Phonétique acoustique [Matériel de cours Service d'Enseignement à Distance]*.
- Astésano, C. (2016). Prosodic characteristics of Reference French. In *Varieties of Spoken French* (Issue January 2016, pp. 68–85). Oxford University Press Oxford. <https://doi.org/10.1093/acprof:oso/9780199573714.003.0006>
- Astésano, C. (2017). *Le statut de l'accent initial dans la phonologie prosodique du français – Enjeux descriptifs et psycholinguistiques* (Vol. 1). Université Toulouse 2.
- Astésano, C. (2019). The prosodic word as the domain of French accentuation -Empirical evidence. *Phonetics and Phonology in Europe, PaPE*, 170–171.
- Astésano, C. (2022). De la supramodalité du rythme : Implications pour la description prosodique, la remédiation linguistique et l'apprentissage des langues. *XXXIVe Journées d'Études Sur La Parole – JEP 2022, June*, 1–14. <https://doi.org/10.21437/jep.2022-1>
- Astésano, C., & Bertrand, R. (2016). Accentuation et niveaux de constitution en français : enjeux phonologiques et psycholinguistiques. *Langue Française, 191*(3), 11.

<https://doi.org/10.3917/lf.191.0011>

Astésano, C., Yagoubi, R. El, & Aguilera, M. (2013). Processing of the Initial Accent by French listeners : investigation of pre-attentional and attentional processing in an oddball paradigm . *Amlap*, 2007, 2013.

Audacity Team. (2021). *Audacity(R): Free Audio Editor and Recorder*.

Augustin. (1878). De la musique. In *Oeuvres complètes de saint Augustin - Tome 3*. Louis Vivès.

Bagou, O., & Frauenfelder, U. H. (2002). Alignement lexical et segmentation de la parole. *Revue Française de Linguistique Appliquée*, VII(1), 67. <https://doi.org/10.3917/rfla.071.0067>

Baqué, L., Barbeau, E., Sahraoui, H., & Nespoulous, J.-L. (2016). Les aphasies : pierres angulaires de la neuro(psycho)linguistique. In S. Pinto & M. Sato (Eds.), *Traité de Neurolinguistique* (pp. 295–313). De Boeck Supérieur.

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(JUN), 1–11. <https://doi.org/10.3389/fnins.2016.00245>

Benveniste, E. (1966). Problèmes de linguistique générale, Tome 1. In *Lingua*. Gallimard. [https://doi.org/10.1016/0024-3841\(67\)90056-3](https://doi.org/10.1016/0024-3841(67)90056-3)

Besson, M., Chobert, J., & Marie, C. (2011). Transfer of training between music and speech: Common processing, attention, and memory. *Frontiers in Psychology*, 2(MAY), 1–12. <https://doi.org/10.3389/fpsyg.2011.00094>

Boersma, P., & Weenink, D. (2022). *Praat: doing phonetics by computer* (6.2.19).

Brady, J. P. (1969). Studies on the metronome effect on stuttering. *Behaviour Research and Therapy*, 7(2), 197–204. [https://doi.org/10.1016/0005-7967\(69\)90033-3](https://doi.org/10.1016/0005-7967(69)90033-3)

Brandt, A., Gebrian, M., & Slevc, L. R. (2012). Music and early language acquisition. *Frontiers in Psychology*, 3(SEP), 1–17. <https://doi.org/10.3389/fpsyg.2012.00327>

Bridges, D., Pitiot, A., MacAskill, M. R., & Peirce, J. W. (2020). The timing mega-study: Comparing a range of experiment generators, both lab-based and online. *PeerJ*, 8, 1–29.

<https://doi.org/10.7717/peerj.9414>

- Buchner, A., Erdfelder, E., Faul, F., & Lang, A. G. (2021). *G * Power 3.1 manual*. https://www.psychologie.hhu.de/fileadmin/redaktion/Fakultaeten/Mathematisch-Naturwissenschaftliche_Fakultaet/Psychologie/AAP/gpower/GPowerManual.pdf
- Budisavljevic, S., Dell'Acqua, F., Djordjilovic, V., Miotto, D., Motta, R., & Castiello, U. (2017). The role of the frontal aslant tract and premotor connections in visually guided hand movements. *NeuroImage*, *146*(April 2016), 419–428. <https://doi.org/10.1016/j.neuroimage.2016.10.051>
- 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. (2020). Regular rhythmic primes boost P600 in grammatical error processing in dyslexic adults and matched controls. *Neuropsychologia*, *138*(December 2019). <https://doi.org/10.1016/j.neuropsychologia.2019.107324>
- Cason, N., Astésano, C., & Schön, D. (2015). 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. (2015). Rhythmic priming enhances speech production abilities: Evidence from prelingually deaf children. *Neuropsychology*, *29*(1), 102–107. <https://doi.org/10.1037/neu0000115>
- Cason, N., Marmursztejn, M., D'Imperio, M., & Schön, D. (2020). Rhythmic Abilities Correlate with L2 Prosody Imitation Abilities in Typologically Different Languages. *Language and Speech*, *63*(1), 149–165. <https://doi.org/10.1177/0023830919826334>
- 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>
- Ceriani, G. (2000). *Du dispositif rythmique - Arguments pour une sémio-physique*. L'Harmattan.
- Chemin, B., Mouraux, A., & Nozaradan, S. (2014). Body Movement Selectively Shapes the

- Neural Representation of Musical Rhythms. *Psychological Science*, 25(12), 2147–2159. <https://doi.org/10.1177/0956797614551161>
- Chen, A. H., Khalid, N. M., & Buari, N. H. (2019). Age factor affects reading acuity and reading speed in attaining text information. *International Journal of Ophthalmology*, 12(7), 1170–1176. <https://doi.org/10.18240/ijo.2019.07.19>
- Chen, J. L., Penhune, V. B., & Zatorre, R. J. (2008). Listening to musical rhythms recruits motor regions of the brain. *Cerebral Cortex*, 18(12), 2844–2854. <https://doi.org/10.1093/cercor/bhn042>
- Chobert, J., François, C., Velay, J. L., & Besson, M. (2014). Twelve months of active musical training in 8-to 10-year-old children enhances the preattentive processing of syllabic duration and voice onset time. *Cerebral Cortex*, 24(4), 956–967. <https://doi.org/10.1093/cercor/bhs377>
- Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (Third Edit). Lawrence Erlbaum Associates.
- Contreras Roa, L., Mairano, P., Lacroix, J., Riegler, R., Moreau, C., & Basirat, A. (2022). Does listening to non-linguistic rhythm impact speech production ? *ESCOP*.
- Corriveau, K. H., & Goswami, U. (2009). Rhythmic motor entrainment in children with speech and language impairments: Tapping to the beat. *Cortex*, 45(1), 119–130. <https://doi.org/10.1016/j.cortex.2007.09.008>
- Corriveau, K. H., Pasquini, E., & Goswami, U. (2007). Basic auditory processing skills and specific language impairment: A new look at an old hypothesis. *Journal of Speech, Language, and Hearing Research*, 50(3), 647–666. [https://doi.org/10.1044/1092-4388\(2007/046\)](https://doi.org/10.1044/1092-4388(2007/046))
- Cutler, A., & Norris, D. (1988). The Role of Strong Syllables in Segmentation for Lexical Access. *Journal of Experimental Psychology: Human Perception and Performance*, 14(1), 113–121. <https://doi.org/10.1037/0096-1523.14.1.113>
- Daigmorte, C., Tallet, J., & Astésano, C. (2022). On the foundations of rhythm-based methods in Speech Therapy. *Speech Prosody* 2022, May, 47–51. <https://doi.org/10.21437/speechprosody.2022-10>

- Dalla Bella, S., Farrugia, N., Benoit, C. E., Bégel, V., Verga, L., Harding, E. 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>
- Delattre, P. (1966). A comparison of syllable length conditioning among languages. *IRAL - International Review of Applied Linguistics in Language Teaching*, *4*(1–4), 183–198. <https://doi.org/10.1515/iral.1966.4.1-4.183>
- Dell, F. (1984). L'accentuation dans les phrases en français. In F. Dell, D. Hirst, & J.-R. Vergnaud (Eds.), *Forme sonore du langage* (pp. 65–122). Hermann.
- Di Cristo, A. (2003). De la métrique et du rythme de la parole ordinaire : l'exemple du français. *Semen*, *16*. <https://doi.org/10.4000/semen.2944>
- Di Cristo, A. (2010). Le pouvoir de la prosodie ou la revanche de Cendrillon. *La Langue et l'être Communiquant. Hommage à Julio Murillo*, 95–113.
- 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. In *Book*. De Gruyter.
- Di Cristo, A., & Hirst, D. (1993). *Rythme syllabique, rythme mélodique et représentation hiérarchique de la prosodie du français*. January.
- Dick, A. S., Garic, D., Graziano, P., & Tremblay, P. (2019). The frontal aslant tract (FAT) and its role in speech, language and executive function. *Cortex*, *111*, 148–163. <https://doi.org/10.1016/j.cortex.2018.10.015>
- Dodane, C., Martel, K., & Vasconcelos, A. N. de. (2018). La prosodie du langage adressé à l'enfant comparée à celle du langage adressé à l'adulte : analyse d'un corpus en français, anglais et japonais. *Cahiers de Praxématique*, *70*. <https://doi.org/10.4000/praxématique.4820>
- Dohen, M., Lœvenbruck, H., Roustan, B., & Vilain, C. (2011). Pointage manuel et vocal: corrélats sensorimoteurs et neurocognitifs, coordination geste manuel / parole. *Faits de Langues*, *37*(1), 97–116. <https://doi.org/10.1163/19589514-037-01-900000006>
- Drake, C. (1998). Psychological Processes Involved in the Temporal Organization of Complex Auditory Sequences: Universal and Acquired Processes. *Music Perception: An*

Interdisciplinary Journal, 16(1), 11–26.

- Drake, C. (2010). Ecouter et jouer de la musique : une fenêtre sur les processus d'organisation temporelle. In B. Lechevalier, H. Platel, & F. Eustache (Eds.), *Le Cerveau musicien*. De Boeck Supérieur.
- Drake, C., Jones, M. R., & Baruch, C. (2000). The development of rhythmic attending in auditory sequences: Attunement, referent period, focal attending. In *Cognition* (Vol. 77, Issue 3). [https://doi.org/10.1016/S0010-0277\(00\)00106-2](https://doi.org/10.1016/S0010-0277(00)00106-2)
- Fadiga, L., Craighero, L., Buccino, G., & Rizzolatti, G. (2002). Speech listening specifically modulates the excitability of tongue muscles: A TMS study. *European Journal of Neuroscience*, 15(2), 399–402. <https://doi.org/10.1046/j.0953-816x.2001.01874.x>
- 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., Müller, T., & Dalla Bella, S. (2015). Non-verbal sensorimotor timing deficits in children and adolescents who stutter. *Frontiers in Psychology*, 6(July), 1–12. <https://doi.org/10.3389/fpsyg.2015.00847>
- Falk, S., Volpi-Moncorger, C., & Dalla Bella, S. (2017). Auditory-motor rhythms and speech processing in French and German listeners. *Frontiers in Psychology*, 8(MAR), 1–14. <https://doi.org/10.3389/fpsyg.2017.00395>
- Fant, G., Kruckenberg, A., & Nord, L. (1991). Durational correlates of stress in Swedish, French and English. *Journal of Phonetics*, 19(3–4), 351–365. [https://doi.org/10.1016/s0095-4470\(19\)30327-4](https://doi.org/10.1016/s0095-4470(19)30327-4)
- Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2), 175–191. <https://doi.org/10.3758/BF03193146>
- 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>. Processing
- Flaunacco, 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).
<https://doi.org/10.1371/journal.pone.0138715>
- Fowler, C. A. (1979). “Perceptual centers” in speech production and perception. *Perception & Psychophysics*, 25(5), 375–388. <https://doi.org/10.3758/BF03199846>
- Fraisse, P. (1966). L’anticipation de stimulus rythmiques. Vitesse d’établissement et précision de la synchronisation. *L’Année Psychologique*, 66(1), 15–36.
- Fraisse, P. (1974). *Psychologie du rythme*. Presses Universitaires de France.
- François, C., Chobert, J., Besson, M., & Schön, D. (2013). Music training for the development of speech segmentation. *Cerebral Cortex*, 23(9), 2038–2043.
<https://doi.org/10.1093/cercor/bhs180>
- François, C., & Schön, D. (2011). Musical expertise boosts implicit learning of both musical and linguistic structures. *Cerebral Cortex*, 21(10), 2357–2365.
<https://doi.org/10.1093/cercor/bhr022>
- Friederici, A. D. (2017). *Language in our brain*. MIT Press.
- Gallucci, M. (2019). *GAMLj: General analyses for linear models. [jamovi module]*.
<https://gamlj.github.io/>
- Garnier, L. (2018). *Quels liens entre accentuation et niveaux de constituance en français? Une analyse perceptive et acoustique*.
- Gatto, R. G., Martin, P. R., Utianski, R. L., Duffy, J. R., Clark, H. M., Botha, H., Machulda, M. M., Josephs, K. A., & Whitwell, J. L. (2024). Diffusion tensor imaging-based multi-fiber tracking reconstructions can regionally differentiate phonetic versus prosodic subtypes of progressive apraxia of speech. *Cortex*, 171, 272–286.
<https://doi.org/10.1016/j.cortex.2023.08.019>
- Gentilucci, M., & Corballis, M. C. (2006). From manual gesture to speech: A gradual transition. *Neuroscience and Biobehavioral Reviews*, 30(7), 949–960.
<https://doi.org/10.1016/j.neubiorev.2006.02.004>
- Gillioz, C., & Zufferey, S. (2020). *Introduction à la linguistique expérimentale*. ISTE Editions.
- Gindre, A.-F., Sahraoui, H., & Astésano, C. (2023). Effet d’amorces rythmiques sur le temps

- d'initiation de la parole chez des adultes tout-venant. In *9èmes Journées de Phonétique Clinique*.
- Goldman, J. P. (2011). EasyAlign: an automatic phonetic alignment tool under Praat. *InterSpeech*.
- Goldman, J. P., Auchlin, A., Roekhaut, S., Simon, A. C., & Avanzi, M. (2010). Prominence perception and accent detection in French. A corpus-based account. *Proceedings of the International Conference on Speech Prosody, May 2014*.
- Gould, L., McKibben, T., Ekstrand, C., Lorentz, E., & Borowsky, R. (2016). The beat goes on: the effect of rhythm on reading aloud. *Language, Cognition and Neuroscience, 31*(2), 236–250. <https://doi.org/10.1080/23273798.2015.1089360>
- Grabski, K., & Sato, M. (2016). Chapitre 6. Percevoir, agir, communiquer : la nature sensorimotrice de la parole. In S. Pinto & M. Sato (Eds.), *Traité de Neurolinguistique* (pp. 83–95). De Boeck Supérieur.
- 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>
- Guiraud, H., Bedoin, N., Krifi-Papoz, S., Herbillon, V., Caillot-Bascoul, A., Gonzalez-Monge, S., & Boulenger, V. (2018). Don't speak too fast! Processing of fast rate speech in children with specific language impairment. *PLoS ONE, 13*(1), 1–23. <https://doi.org/10.1371/journal.pone.0191808>
- 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*(JAN). <https://doi.org/10.3389/fpsyg.2016.00026>
- Hagoort, P., & Levelt, W. J. M. (2009). The speaking brain. *Science, 326*(5951), 372–373. <https://doi.org/10.1126/science.1181675>
- Hickok, G., & Poeppel, D. (2004). Dorsal and ventral streams: A framework for understanding aspects of the functional anatomy of language. *Cognition, 92*(1–2), 67–99. <https://doi.org/10.1016/j.cognition.2003.10.011>
- Hickok, G., & Poeppel, D. (2007). The cortical organization of speech understanding. *Nature, 8*(May), 393–402. www.nature.com/reviews/neuro <https://www-nature-com.ezp->

prod1.hul.harvard.edu/articles/nrn2113.pdf

- Indefrey, P., & Levelt, W. J. M. (2004). The spatial and temporal signatures of word production components. *Cognition*, 92(1–2), 101–144. <https://doi.org/10.1016/j.cognition.2002.06.001>
- Iverson, J. M., & Goldin-Meadow, S. (1998). Why people gesture when they speak. *Nature*, 396(6708), 228. <https://doi.org/10.1038/24300>
- Iverson, J. M., & Thelen, E. (1999). Hand, mouth and brain: The dynamic emergence of speech and gesture. *Journal of Consciousness Studies*, 6(11–12), 19–40.
- Jankowski, L., Astésano, C., & Di Cristo, A. (1999). the Initial Rhythmic Accent in French : Acoustic Data and Perceptual Investigation. *14th International Congress of Phonetic Sciences*, 257–260.
- Jungers, M. K., Hupp, J. M., & Dickerson, S. D. (2016). Language priming by music and speech: Evidence of a shared processing mechanism. *Music Perception*, 34(1), 33–39. <https://doi.org/10.1525/mp.2016.34.1.33>
- Kemerdere, R., de Champfleury, N. M., Deverdun, J., Cochereau, J., Moritz-Gasser, S., Herbet, G., & Duffau, H. (2016). Role of the left frontal aslant tract in stuttering: a brain stimulation and tractographic study. *Journal of Neurology*, 263(1), 157–167. <https://doi.org/10.1007/s00415-015-7949-3>
- Kirschner, S., & Tomasello, M. (2009). Joint drumming: Social context facilitates synchronization in preschool children. *Journal of Experimental Child Psychology*, 102(3), 299–314. <https://doi.org/10.1016/j.jecp.2008.07.005>
- 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(1), 62–68. <https://doi.org/10.1111/nyas.12657>
- Kronfeld-Duenias, V., Amir, O., Ezrati-Vinacour, R., Civier, O., & Ben-Shachar, M. (2016). The frontal aslant tract underlies speech fluency in persistent developmental stuttering. *Brain Structure and Function*, 221(1), 365–381. <https://doi.org/10.1007/s00429-014-0912-8>
- 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). <https://doi.org/10.1002/wcs.1528>
- Large, E. W., Fink, P., & Kelso, J. A. S. (2002). Tracking simple and complex sequences. *Psychological Research*, 66(1), 3–17. <https://doi.org/10.1007/s004260100069>
- Large, E. W., & Jones, M. R. (1999). The dynamics of attending: How people track time-varying events. *Psychological Review*, 106(1), 119–159.
- Lê, M., Quémart, P., Potocki, A., Gimenes, M., Chesnet, D., & Lambert, E. (2020). Rhythm in the blood: The influence of rhythm skills on literacy development in third graders. *Journal of Experimental Child Psychology*, 198. <https://doi.org/10.1016/j.jecp.2020.104880>
- Lenc, T., Merchant, H., Keller, P. E., Honing, H., Varlet, M., & Nozaradan, S. (2021). Mapping between sound, brain and behaviour: Four-level framework for understanding rhythm processing in humans and non-human primates. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 376(1835). <https://doi.org/10.1098/rstb.2020.0325>
- Levelt, W. J. M. (1999). Models of word production. *Trends in Cognitive Sciences*, 3(6), 223–232.
- Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, 22(1), 1–75. <https://doi.org/10.1017/S0140525X99001776>
- Lieberman, A. M., & Mattingly, I. G. (1985). The motor theory of speech perception revised. *Cognition*, 21, 1–36.
- LimeSurvey Project Team. (n.d.). *LimeSurvey: An Open Source survey tool*. <https://www.limesurvey.org>
- Lœvenbruck, H. (2019). *Loquor, ergo communico-cogito-sum*. Université Grenoble Alpes.
- Lœvenbruck, H., Dohen, M., & Vilain, C. (2009). Pointing is “special.” *Some Aspects of Speech and the Brain*, May 2014, 211–258.
- MacNeilage, P. F. (1998). The frame/content theory of evolution of speech production. *Behavioral and Brain Sciences*, 21(4), 499–546. <https://doi.org/10.1017/S0140525X98001265>
- Magne, C. L., Astésano, C., Aramaki, M., Ystad, S., Kronland-Martinet, R., & Besson, M. (2007). Influence of syllabic lengthening on semantic processing in spoken French:

- Behavioral and electrophysiological evidence. *Cerebral Cortex*, 17(11), 2659–2668. <https://doi.org/10.1093/cercor/bhl174>
- McAuley, J. D., Jones, M. R., Holub, S., Johnston, H. M., & Miller, N. S. (2006). The time of our lives: Life span development of timing and event tracking. *Journal of Experimental Psychology: General*, 135(3), 348–367. <https://doi.org/10.1037/0096-3445.135.3.348>
- Méndez, R., & Astésano, C. (2017). Perception of the Downstepped Final Accent in French. *Phonetics and Phonology in Europe 2017, June*, 102–103. <https://doi.org/10.13140/RG.2.2.28176.20488>
- Microsoft Corporation. (2021). *Microsoft Excel*.
- Miura, A., Kudo, K., Ohtsuki, T., & Kanehisa, H. (2011). Coordination modes in sensorimotor synchronization of whole-body movement: A study of street dancers and non-dancers. *Human Movement Science*, 30(6), 1260–1271. <https://doi.org/10.1016/j.humov.2010.08.006>
- Monfrais-Pfauwadel, M.-C., & Bijleveld, H. (2016). Le bégaiement, les bégaiements. In *Traité de neurolinguistique* (pp. 265–277). De Boeck Supérieur.
- Monier, F., & Droit-Volet, S. (2019). Development of sensorimotor synchronization abilities: Motor and cognitive components. *Child Neuropsychology*, 25(8), 1043–1062. <https://doi.org/10.1080/09297049.2019.1569607>
- Morillon, B., Schroeder, C. E., & Wyart, V. (2014). Motor contributions to the temporal precision of auditory attention. *Nature Communications*, 5. <https://doi.org/10.1038/ncomms6255>
- Moritz, C., Yampolsky, S., Papadelis, G., Thomson, J. M., & Wolf, M. (2013). Links between early rhythm skills, musical training, and phonological awareness. *Reading and Writing*, 26(5), 739–769. <https://doi.org/10.1007/s11145-012-9389-0>
- 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>
- Moussard, A., Rochette, F., & Bigand, E. (2012). La musique comme outil de stimulation cognitive. *L'Année Psychologique*, 112(03), 499–542. <https://doi.org/10.4074/s0003503312003077>

- Nespoulous, J.-L. (2016). La neuro(psycho)linguistique : Emergence, évolution et perspectives d'une discipline à la croisée des chemins. In *Traité de Neurolinguistique* (pp. 1–12).
- Norton, A. C., Zipse, L., Marchina, S., & Schlaug, G. (2009). Melodic Intonation Therapy: Shared Insights on How it is Done and Why it Might Help. *Stroke*, *1169*, 431–436. <https://doi.org/10.1111/j.1749-6632.2009.04859.x>.
- Ozernov-Palchik, O., Wolf, M., & Patel, A. D. (2018). Relationships between early literacy and nonlinguistic rhythmic processes in kindergarteners. *Journal of Experimental Child Psychology*, *167*(December), 354–368. <https://doi.org/10.1016/j.jecp.2017.11.009>
- Parbery-Clark, A., Skoe, E., & Kraus, N. (2009). Musical experience limits the degradative effects of background noise on the neural processing of sound. *Journal of Neuroscience*, *29*(45), 14100–14107. <https://doi.org/10.1523/JNEUROSCI.3256-09.2009>
- Parbery-Clark, A., Tierney, A. T., Strait, D. L., & Kraus, N. (2012). Musicians have fine-tuned neural distinction of speech syllables. *Neuroscience*, *219*, 111–119. <https://doi.org/10.1016/j.neuroscience.2012.05.042>
- Patel, A. D. (2011). Why would musical training benefit the neural encoding of speech? The OPERA hypothesis. *Frontiers in Psychology*, *2*(JUN), 1–14. <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>
- 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*(MAY), 1–14. <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*(SEP), 1–17. <https://doi.org/10.3389/fpsyg.2012.00320>
- Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195–203. <https://doi.org/10.3758/s13428-018-01193-y>
- Phillips-silver, J., & Trainor, L. J. (2005). Feeling the Beat : Movement Influences Infant

- Rhythm Perception. *Science*, 308, 1430. <https://doi.org/10.1126/science.1110922>
- Phillips-Silver, J., & Trainor, L. J. (2007). Hearing what the body feels: Auditory encoding of rhythmic movement. *Cognition*, 105(3), 533–546. <https://doi.org/10.1016/j.cognition.2006.11.006>
- Piaget, J. (1942). Les trois structures fondamentales de la vie psychique: rythme, regulation et groupement. *Revue Suisse de Psychologie*, 1(1), 9–21.
- Pitt, M. A., & Samuel, A. G. (1990). The Use of Rhythm in Attending to Speech. *Journal of Experimental Psychology: Human Perception and Performance*, 16(3), 564–573. <https://doi.org/10.1037/0096-1523.16.3.564>
- Port, R. F., Cummins, F., & Gasser, M. (1995). A Dynamic Approach to Rhythm in Language: Toward a Temporal Phonology. *Computation and Language E-Print Archive*, 1–31. <http://arxiv.org/abs/cmp-lg/9508007>
- Postma, A. (2000). Detection of errors during speech production: A review of speech monitoring models. *Cognition*, 77(2), 97–132. [https://doi.org/10.1016/S0010-0277\(00\)00090-1](https://doi.org/10.1016/S0010-0277(00)00090-1)
- Povel, D.-J., & Essens, P. J. (1985). Perception of Temporal Patterns. *Music Perception: An Interdisciplinary Journal*, 2(4), 411–440.
- Provasi, J. (2016). Comment le rythme vient aux Bébés? *Spirale*, 76(4), 50–63. <https://doi.org/10.3917/spi.076.0050>
- Provasi, J., & Bobin-Bègue, A. (2003). Spontaneous motor tempo and rhythmical synchronisation in 21/2- and 4-year-old children. *International Journal of Behavioral Development*, 27(3), 220–231. <https://doi.org/10.1080/01650250244000290>
- 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>
- R Core Team. (2021). *R: A Language and Environment for Statistical Computing*.
- Ramus, F., Nespors, M., & Mehler, J. (1999). Correlates of linguistic rhythm in the speech signal. *Cognition*, 75(1), 265–292. [https://doi.org/10.1016/s0010-0277\(00\)00101-3](https://doi.org/10.1016/s0010-0277(00)00101-3)

- Repp, B. H. (2003). Rate limits in sensorimotor synchronization with auditory and visual sequences: the synchronization threshold and the benefits and costs of interval subdivision. *Journal of Motor Behavior*, 35(4), 355–370.
- Repp, B. H. (2005). Sensorimotor synchronization: A review of the tapping literature. *Psychonomic Bulletin and Review*, 12(6), 969–992. <https://doi.org/10.3758/BF03206433>
- Roodenrys, S., & Stokes, J. (2001). Serial recall and nonword repetition in reading disabled children. *Reading and Writing*, 14(5–6), 379–394. <https://doi.org/10.1023/a:1011123406884>
- Sahraoui, H. (2015). Ce que nous apprennent les troubles du langage. In J.-F. Dortier (Ed.), *Les clés du langage* (pp. 42–50). Editions Sciences Humaines. <https://doi.org/sh.dorti.2015.01.0042>
- Särkämö, T., Pihko, E., Laitinen, S., Forsblom, A., Soinila, S., Mikkonen, M., Autti, T., Silvennoinen, H. M., Erkkilä, J., Laine, M., Peretz, I., Hietanen, M., & Tervaniemi, M. (2010). Music and speech listening enhance the recovery of early sensory processing after stroke. *Journal of Cognitive Neuroscience*, 22(12), 2716–2727. <https://doi.org/10.1162/jocn.2009.21376>
- Sauvanet, P. (2000). *Le Rythme et la raison - Tome 1 : Rythmologiques*. Kimé.
- Schön, D., Hidalgo, C., & Tillmann, B. (2018). Stimuler le langage par la musique. In E. Bigand (Ed.), *Les bienfaits de la musique sur le cerveau*. Belin.
- Schön, D., Magne, C. L., & Besson, M. (2004). The music of speech: Music training facilitates pitch processing in both music and language. *Psychophysiology*, 41(3), 341–349. <https://doi.org/10.1111/1469-8986.00172.x>
- Schwartz, J., Basirat, A., Ménard, L., & Sato, M. (2012). The Perception-for-Action-Control Theory (PACT): A perceptuo-motor theory of speech perception. *Journal of Neurolinguistics*, 25(5), 336–354. <https://doi.org/10.1016/j.jneuroling.2009.12.004>
- Shao, Z., & Meyer, A. S. (2018). Word Priming and Interference Paradigms. In A. M. B. de Groot & P. Hagoort (Eds.), *Research Methods in Psycholinguistics and the Neurobiology of Language : A Practical Guide*. Wiley Blackwell.
- Shields, J. L., McHugh, A., & Martin, J. G. (1974). Reaction time to phoneme targets as a function of rhythmic cues in continuous speech. *Journal of Experimental Psychology*,

- 102(2), 250–255. <https://doi.org/10.1037/h0035855>
- Slater, J., & Kraus, N. (2016). The role of rhythm in perceiving speech in noise: a comparison of percussionists, vocalists and non-musicians. *Physiology & Behavior*, 176(1), 139–148. <https://doi.org/10.1016/j.physbeh.2017.03.040>
- Snyder, J. S., & Large, E. W. (2005). Gamma-band activity reflects the metric structure of rhythmic tone sequences. *Cognitive Brain Research*, 24(1), 117–126. <https://doi.org/10.1016/j.cogbrainres.2004.12.014>
- Solier, C., Soulier, L., & Ezzedine, N. (2023). *Introduction aux statistiques en sciences du langage: Traitement et analyse de données avec R*. Dunod. <https://books.google.fr/books?id=lyHbEAAAQBAJ>
- Sowiński, J., & Dalla Bella, S. (2013). Poor Synchronization to the Beat May Result From Deficient Auditory-Motor Mapping. *Neuropsychologia*, 51(10), 1952–1963. <https://doi.org/10.1016/j.neuropsychologia.2013.06.027>
- 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 and Phonetics*, 30(1), 68–85. <https://doi.org/10.3109/02699206.2015.1115129>
- Speer, P., & Wilshire, C. E. (2014). What's in a sentence? The crucial role of lexical content in sentence production in nonfluent aphasia. *Cognitive Neuropsychology*, 1–37. <https://doi.org/10.1080/02643294.2013.876398>
- 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>
- Su, Y.-H., & Pöppel, E. (2012). Body movement enhances the extraction of temporal structures in auditory sequences. *Psychological Research*, 76(3), 373–382. <https://doi.org/10.1007/s00426-011-0346-3>
- Sun, Y., & Peperkamp, S. (2016). The role of speech production in phonological decoding during visual word recognition: evidence from phonotactic repair. *Language, Cognition and Neuroscience*, 31(3), 391–403. <https://doi.org/10.1080/23273798.2015.1100316>
- Te Rietmolen, N. (2019). *Neural signature of metrical stress processing in French* [thèse de

- doctorat]. Université Toulouse 2.
- Te Rietmolen, N., El Yagoubi, R., Espesser, R., Magnen, C., & Astésano, C. (2016). Investigating the phonological status of the initial accent in French: An event-related potentials study. *Proceedings of the International Conference on Speech Prosody, 2016-Janua*, 1181–1185. <https://doi.org/10.21437/speechprosody.2016-243>
- Thaut, M. H., & Hoemberg, V. (2019). *Manuel clinique de rééducation par la musique*. De Boeck Supérieur.
- The jamovi Project. (2023). *jamovi* (2.4).
- Thompson, E. C., White-Schwoch, T., Tierney, A. T., & Kraus, N. (2015). Beat synchronization across the lifespan: Intersection of development and musical experience. *PLoS ONE*, *10*(6), 1–13. <https://doi.org/10.1371/journal.pone.0128839>
- Thomson, J. M., Fryer, B., Maltby, J., & Goswami, U. (2006). Auditory and motor rhythm awareness in adults with dyslexia. *Journal of Research in Reading*, *29*(3), 334–348. <https://doi.org/10.1111/j.1467-9817.2006.00312.x>
- Tierney, A., & Kraus, N. (2013). The ability to move to a beat is linked to the consistency of neural responses to sound. *Journal of Neuroscience*, *33*(38), 14981–14988. <https://doi.org/10.1523/JNEUROSCI.0612-13.2013>
- Tierney, A. T., & Kraus, N. (2013). The ability to tap to a beat relates to cognitive, linguistic, and perceptual skills. *Brain and Language*, *124*, 225–231. <https://doi.org/10.1016/j.bandl.2009.07.002>
- Tourville, J. A., & Guenther, F. H. (2011). The DIVA model: A neural theory of speech acquisition and production. *Language and Cognitive Processes*, *26*(7), 1–27. <https://doi.org/10.1080/01690960903498424>
- Trayssac, E. (2021). *Développement des capacités rythmiques chez les adolescents porteurs de déficience intellectuelle légère idiopathique* [mémoire d'orthophonie]. Université Toulouse 3.
- Turk, A., & Shattuck-Hufnagel, S. (2014). Timing in talking: What is it used for, and how is it controlled? *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(1658). <https://doi.org/10.1098/rstb.2013.0395>

- Vaissière, J. (2020). Quelques aspects de la perception de la parole. In J. Vaissière (Ed.), *La phonétique* (pp. 79–95). Presses Universitaires de France.
- van de Vorst, R., & Gracco, V. L. (2017). Atypical non-verbal sensorimotor synchronization in adults who stutter may be modulated by auditory feedback. *Journal of Fluency Disorders*, 53(November 2016), 14–25. <https://doi.org/10.1016/j.jfludis.2017.05.004>
- Vaysse, R. (2023). *Caractérisation automatique du rythme de la parole : application aux cancers des voies aéro-digestives supérieures et à la maladie de Parkinson*. Université Toulouse 3.
- Véron-Delor, L., Velay, J., Braibant, I., & Danna, J. (2017). Qu’apporte la musique à l’apprentissage de l’écriture ? Étude de cas auprès d’un enfant en grande difficulté d’écriture. *A.N.A.E. Approche Neuropsychologique Des Apprentissages Chez l’enfant*.
- Vilain, A., Dole, M., Lœvenbruck, H., Pascalis, O., & Schwartz, J. L. (2019). The role of production abilities in the perception of consonant category in infants. *Developmental Science*, 22(6). <https://doi.org/10.1111/desc.12830>
- Watkins, K. E., Strafella, A. P., & Paus, T. (2003). Seeing and hearing speech excites the motor system involved in speech production. *Neuropsychologia*, 41(8), 989–994. [https://doi.org/10.1016/S0028-3932\(02\)00316-0](https://doi.org/10.1016/S0028-3932(02)00316-0)
- Wu, V. X., Chi, Y., Lee, J. K., Goh, H. S., Chen, D. Y. M., Haugan, G., Chao, F. F. T., & Klainin-Yobas, P. (2021). The effect of dance interventions on cognition, neuroplasticity, physical function, depression, and quality of life for older adults with mild cognitive impairment: A systematic review and meta-analysis. *International Journal of Nursing Studies*, 122, 104025. <https://doi.org/10.1016/j.ijnurstu.2021.104025>
- Zanto, T. P., Giannakopoulou, A., Gallen, C. L., Ostrand, A. E., Younger, J. W., Anguera-Singla, R., Anguera, J. A., & Gazzaley, A. (2024). Digital rhythm training improves reading fluency in children. *Developmental Science*, 27(3). <https://doi.org/10.1111/desc.13473>
- 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>
- Ziegler, W., Lehner, K., Pfab, J., & Aichert, I. (2021). The nonlinear gestural model of speech apraxia: clinical implications and applications. *Aphasiology*, 35(4), 462–484.

<https://doi.org/10.1080/02687038.2020.1727839>

Zuur, A. F., Ieno, E. N., Walker, N., Saveliev, A. A., & Smith, G. M. (2009). Mixed effects models and extensions in ecology with R. In *Media*. Springer New York.
<https://doi.org/10.1007/978-0-387-87458-6>

ANNEXES – PARTIE 1

Annexe 1 Notice d'information et consentement éclairé

NOTICE D'INFORMATION ET CONSENTEMENT ECLAIRE

Titre du projet : Rythme(s) langagier, musical et moteur : quels transferts possibles ?

Chercheure titulaire responsable scientifique du projet :

ASTESANO Corine (co-directrice de thèse)

astesano@univ-tlse2.fr

05.61.50.24.69

Maître de conférences HDR

Laboratoire de NeuroPsychoLinguistique (LNPL)

Université Toulouse - Jean Jaurès - Maison de la Recherche - Bureau E1.15 - 5 allées

Antonio Machado 31058 TOULOUSE Cedex 9

Autres chercheurs participant au projet :

SAHRAOUI Halima (co-directrice de thèse)

halima.sahraoui@univ-tlse2.fr

05.61.50.24.72

Maître de conférences

Laboratoire de NeuroPsychoLinguistique (LNPL)

Université Toulouse - Jean Jaurès - Maison de la Recherche - Bureau E1.26 - 5 allées

Antonio Machado 31058 TOULOUSE Cedex 9

GINDRE Anne-Flore (doctorante)

anne-flore.gindre@univ-tlse2.fr

05.61.50.35.98

Laboratoire de NeuroPsychoLinguistique (LNPL)

Université Toulouse - Jean Jaurès - Maison de la Recherche - Bureau E1.26 - 5 allées

Antonio Machado 31058 TOULOUSE Cedex 9

Lieu de recherche :

Université Toulouse - Jean Jaurès, Maison de la Recherche, plateforme CCU.

But du projet de recherche :

L'objectif de ce projet de recherche de thèse est d'évaluer les liens entre rythmes langagier, musical et moteur (gestuel) auprès de la population générale. Les résultats pourront mener à des applications thérapeutiques de rééducation de la fluence de la parole (aphasie, bégaiement, troubles neurodéveloppementaux, etc.).

Vous pouvez participer à cette étude si vous remplissez tous les critères suivants :

- vous êtes âgé-e d'au moins 18 ans,
- le français est votre langue maternelle,
- votre vision est normale ou corrigée,
- vous ne présentez pas de trouble de l'audition, de la motricité ou du langage,
- vous n'êtes pas professionnel-le dans le domaine de la musique ou de la danse et
- vous êtes volontaire pour participer et signez le présent formulaire de consentement de participation.

Ce que l'on attend de vous (méthodologie) :

Si vous consentez à prendre part à cette étude, vous participerez à deux activités expérimentales. Lors de la première activité, nous prendrons la mesure de votre tempo naturel et de votre synchronisation à un métronome (environ 5 minutes). Lors de la seconde activité, vous devrez lire de courtes phrases à haute voix (environ 35 minutes). Nous enregistrerons votre mouvement de la main et votre voix sur les phrases lues. Plusieurs pauses seront proposées au cours de la passation.

Vos droits de vous retirer de la recherche en tout temps :

Vous confirmez participer à cette expérience de votre pleine volonté et êtes informé-e que vous pouvez à tout moment cesser votre participation sans avoir à le justifier. Cette décision de participer ou de vous retirer n'aura aucun effet sur votre parcours étudiant (si concerné), votre statut, vos relations futures avec le laboratoire de NeuroPsychoLinguistique (LNPL) ou l'Université Toulouse - Jean Jaurès.

Vos droits à la confidentialité et au respect de la vie privée :

Les informations recueillies par le laboratoire LNPL vous concernant ont fait l'objet d'une déclaration en juin 2022 auprès du service de Protection des données de l'Université Toulouse - Jean Jaurès, 5 allées Antonio Machado, 31058 Toulouse Cedex 9, dpo@univ-tlse2.fr (fiche de traitement R-202205161732).

La base légale de cette déclaration s'appuie sur votre consentement.

Les données obtenues seront traitées avec la plus grande confidentialité. Aucun renseignement ne sera dévoilé qui puisse révéler votre identité. Toutes les données seront gardées dans un endroit sécurisé et seules les chercheuses mentionnées ci-après y auront accès : Corine Astésano, Halima Sahraoui et Anne-Flore Gindre.

Vos données expérimentales et personnelles seront gardées en France. Elles seront conservées en base active jusqu'à 2 ans après la dernière publication sur cette recherche. Passée cette échéance, les données identifiantes seront supprimées et les données expérimentales non identifiantes seront archivées dans un endroit sécurisé pour une durée de 10 ans.

Conformément aux dispositions de la loi Informatique et Libertés, vous disposez de droits concernant le traitement de vos données. Après la fin de l'entretien, vos droits à la rectification ou à la suppression de vos informations pourront s'appliquer à condition de fournir le numéro aléatoire qui vous est attribué. Vous pouvez exercer vos droits en vous adressant à la responsable scientifique du projet, Corine Astésano (corine.astesano@univ-tlse2.fr). Si vous estimez, après nous avoir contactées, que vos droits sur vos données ne sont pas respectés, vous pouvez adresser une réclamation (plainte) à la CNIL (<https://www.cnil.fr/fr/adresser-une-plainte>).

Bénéfices :

L'étude n'a pas de bénéfices immédiats pour votre santé physique et mentale ou votre vie sociale.

Les résultats de cette étude permettront de proposer des applications thérapeutiques pour la rééducation de la fluence de la parole (aphasie, bégaiement, pathologies neurodéveloppementales, etc.).

Risques possibles :

Cette étude a reçu l'avis favorable du Comité d'Ethique de la Recherche de Toulouse en juillet 2022 (n° de dossier 2022_545).

Votre participation à l'étude ne présente aucun inconfort ou risque pouvant porter atteinte à votre santé (physique, psychologique ou morale) autres que ceux de la vie quotidienne.

Risques pour la santé liés au contexte sanitaire (COVID-19) : le protocole expérimental nécessite votre venue en personne à l'Université Toulouse - Jean Jaurès. Les recommandations sanitaires mises en place par l'université pour le personnel et les étudiant·e·s sont respectées lors des passations. Des précautions sont prises sur place (désinfection du matériel entre les participant·e·s) afin de limiter les risques de contamination, en fonction des recommandations en vigueur dans l'établissement.

Diffusion :

Les résultats de cette recherche seront diffusés dans des colloques et des articles de revue scientifique.

Vos droits de poser des questions en tout temps :

Vous pouvez poser des questions au sujet de la recherche en tout temps par courrier électronique auprès des chercheurs associés au projet : Corine Astésano (corine.astesano@univ-tlse2.fr), Halima Sahraoui (halima.sahraoui@univ-tlse2.fr) et Anne-Flore Gindre (anne-flore.gindre@univ-tlse2.fr).

Consentement à la participation :

En signant le présent formulaire de consentement, vous certifiez que vous avez lu et compris les renseignements ci-dessus, qu'on a répondu à vos questions de façon satisfaisante et qu'on vous a avisé·e que vous étiez libre d'annuler votre consentement ou de vous retirer de cette recherche en tout temps, sans préjudice.

Votre collaboration est précieuse pour la réalisation de la recherche. Nous vous remercions pour le temps et l'attention que vous consacrez à votre participation.

A remplir par le participant :

- J'accepte que ma voix soit enregistrée.
- J'ai lu et compris les renseignements ci-dessus et j'accepte de plein gré de participer à cette recherche.

Nom, Prénom – Date – Signature

Un exemplaire de ce document vous est remis, un autre exemplaire est conservé dans le dossier.

Annexe 2 Questionnaire socio-démographique

Rythme(s) langagier, musical et moteur : quels transferts possibles ?

Questionnaire sociodémographique : informations générales

1. Âge : _____ ans

2. Sexe :

- Homme
 Femme
 Autre

3. Latéralité :

- Droitier·e
 Gaucher·e
 Ambidextre

4. Plus haut diplôme obtenu :

- Bac
 Bac+3
 Bac+5
 Bac+8
 Autre (précisez) :

5. a. Activité principale (plusieurs réponses possibles) :

- Etudiant·e
 Salarié·e
 Doctorant·e
 Retraité·e
 Sans activité
 Autre (précisez) :

b. Domaine d'activité/étude/métier :

6. Langue(s) parlée(s) :

- Langue(s) maternelle(s) :
- Niveau avancé (C1/C2 du CECRL) :
- Niveau intermédiaire (B1/B2 du CECRL) :
- Niveau débutant (A1/A2 du CECRL) :

7. Dans quel département, région et/ou quel pays avez-vous vécu le plus longtemps ?

8. Faites-vous ou avez-vous fait de la musique ?

- Oui*
 Non

*Si oui, de quel(s) instrument(s) ? Précisez à quelle fréquence et depuis ou pendant combien d'années.

9. Faites-vous ou avez-vous fait de la danse ?

- Oui*
 Non

*Si oui, à quelle fréquence et depuis ou pendant combien d'années ?

10. Jouez-vous ou avez-vous joué aux jeux vidéo (sur console ou ordinateur) ?

- Oui*
 Non

* Si oui, à quelle fréquence et depuis ou pendant combien d'années ?

11. Quel est votre numéro aléatoire ?

Annexe 3 Informations socio-démographiques des participants

Tableau 30 :

Description socio-démographique (âge, années pratique musique et danse)

Statistiques descriptives

| | N | Moyenne | Médiane | Ecart- type | Minimum | Maximum | Shapiro-Wilk | |
|-----------------------|----|---------|---------|----------------|---------|---------|--------------|--------|
| | | | | | | | W | p |
| Age | 40 | 31.60 | 29.000 | 9.38 | 18 | 56 | 0.893 | 0.001 |
| Mus_annees_pratique | 40 | 4.20 | 2.000 | 4.55 | 0 | 15 | 0.851 | < .001 |
| Danse_annees_pratique | 40 | 3.10 | 0.500 | 5.47 | 0 | 21 | 0.623 | < .001 |

Tableau 31 :

Description socio-démographique (sexe)

Fréquences de Sexe

| Sexe | Quantités | % du Total | % cumulés |
|------|-----------|------------|-----------|
| F | 20 | 50.0 % | 50.0 % |
| H | 20 | 50.0 % | 100.0 % |

Tableau 32 :

Description socio-démographique (musique)

Fréquences de Musique

| Musique | Quantités | % du Total | % cumulés |
|---------|-----------|------------|-----------|
| Non | 13 | 32.5 % | 32.5 % |
| Oui | 27 | 67.5 % | 100.0 % |

Tableau 33 :
Description socio-démographique (instrument)

Fréquences de Mus_instrument

| Mus_instrument | Quantités | % du Total | % cumulés |
|----------------|-----------|------------|-----------|
| flûte | 3 | 7.5 % | 7.5 % |
| guitare | 5 | 12.5 % | 20.0 % |
| percussions | 2 | 5.0 % | 25.0 % |
| piano | 6 | 15.0 % | 40.0 % |
| plusieurs | 10 | 25.0 % | 65.0 % |
| violon | 1 | 2.5 % | 67.5 % |
| 0 | 13 | 32.5 % | 100.0 % |

Tableau 34 :
Description socio-démographique (danse)

Fréquences de Danse

| Danse | Quantités | % du Total | % cumulés |
|-------|-----------|------------|-----------|
| Non | 20 | 50.0 % | 50.0 % |
| Oui | 20 | 50.0 % | 100.0 % |

Tableau 35 :
Description socio-démographique (jeux vidéo)

Fréquences de Jeux_video

| Jeux_video | Quantités | % du Total | % cumulés |
|------------|-----------|------------|-----------|
| Non | 12 | 30.0 % | 30.0 % |
| Oui | 28 | 70.0 % | 100.0 % |

Annexe 4 Listes pseudo-aléatoires de l'Activité 2

Tableau 36 :

Listes pseudo-aléatoires de passation (Activité 2)

| participant | ordre global | participant | ordre global |
|-------------|-----------------|-------------|-----------------|
| P01 | SM 123 - AM 123 | Z37 | AM 123 - SM 123 |
| Q02 | SM 132 - AM 123 | A38 | AM 132 - SM 123 |
| R03 | SM 213 - AM 123 | B39 | AM 213 - SM 123 |
| S04 | SM 231 - AM 123 | C40 | AM 231 - SM 123 |
| T05 | SM 312 - AM 123 | D41 | AM 312 - SM 123 |
| U06 | SM 321 - AM 123 | E42 | AM 321 - SM 123 |
| V07 | SM 123 - AM 132 | F43 | AM 123 - SM 132 |
| W08 | SM 132 - AM 132 | G44 | AM 132 - SM 132 |
| X09 | SM 213 - AM 132 | H45 | AM 213 - SM 132 |
| Y10 | SM 231 - AM 132 | I46 | AM 231 - SM 132 |
| Z11 | SM 312 - AM 132 | J47 | AM 312 - SM 132 |
| A12 | SM 321 - AM 132 | K48 | AM 321 - SM 132 |
| B13 | SM 123 - AM 213 | L49 | AM 123 - SM 213 |
| C14 | SM 132 - AM 213 | M50 | AM 132 - SM 213 |
| D15 | SM 213 - AM 213 | N51 | AM 213 - SM 213 |
| E16 | SM 231 - AM 213 | O52 | AM 231 - SM 213 |
| F17 | SM 312 - AM 213 | P53 | AM 312 - SM 213 |
| G18 | SM 321 - AM 213 | Q54 | AM 321 - SM 213 |
| H19 | SM 123 - AM 231 | R55 | AM 123 - SM 231 |
| I20 | SM 132 - AM 231 | S56 | AM 132 - SM 231 |
| J21 | SM 213 - AM 231 | T57 | AM 213 - SM 231 |
| K22 | SM 231 - AM 231 | U58 | AM 231 - SM 231 |
| L23 | SM 312 - AM 231 | V59 | AM 312 - SM 231 |
| M24 | SM 321 - AM 231 | W60 | AM 321 - SM 231 |
| N25 | SM 123 - AM 312 | X61 | AM 123 - SM 312 |
| O26 | SM 132 - AM 312 | Y62 | AM 132 - SM 312 |
| P27 | SM 213 - AM 312 | Z63 | AM 213 - SM 312 |
| Q28 | SM 231 - AM 312 | A64 | AM 231 - SM 312 |
| R29 | SM 312 - AM 312 | B65 | AM 312 - SM 312 |
| S30 | SM 321 - AM 312 | C66 | AM 321 - SM 312 |
| T31 | SM 123 - AM 321 | D67 | AM 123 - SM 321 |
| U32 | SM 132 - AM 321 | E68 | AM 132 - SM 321 |
| V33 | SM 213 - AM 321 | F69 | AM 213 - SM 321 |
| W34 | SM 231 - AM 321 | G70 | AM 231 - SM 321 |
| X35 | SM 312 - AM 321 | H71 | AM 312 - SM 321 |
| Y36 | SM 321 - AM 321 | I72 | AM 321 - SM 321 |

Annexe 5 Matériel expérimental de l'Activité 2

Tableau 37 :

Syntagmes (amorces et cibles) utilisés dans l'Activité 2

| Rythme 1 x x x X x X | Rythme 2 x x X x x X | Rythme 3 x x X x x x X | Rythme 4 x x x X x x X |
|-----------------------------|--------------------------|-------------------------------------|------------------------------|
| Le scandaleux sénat | Elle prépare le buffet | Il prescrit le bon cachet | Chorégrapheur le ballet |
| C'est le meilleur moka | Il atteint la butée | Il adore son frère cadet | Il a plié le bambou |
| C'est un petit dégât | Il a mis son ciré | La voiture a un capot | Il dressera le basset |
| Il vend des vrais bolets | La cuisson du civet | Il habite dans la cité | La religieuse au café |
| Elle a un grand chalet | Relever un défi | J'ai besoin d'un long congé | Les électeurs du comté |
| C'est un très bon filet | Ils convoquent les jurés | Cette affaire est un gâchis | Nous recueillons la couvée |
| Un succulent jarret | Attache bien ton lacet | L'art martial comme le judo | Elle se comporte en diva |
| Ce sont des vieux volets | Je cueillais des lilas | Il repart sur sa lancée | Elle contrôlait la durée |
| Un magnifique dîner | Il protège les manchots | Il recueille un sans-logis | Ils ont volé des lingots |
| L'impopulaire bûcher | Il a pris le maquis | Attention à la marée | J'ai secouru le minet |
| L'éblouissante chaussée | Il chassait le mulot | Je prépare une bonne potée | Il s'est fait mal au mollet |
| Ce n'est qu'un faux guichet | La culture des navets | On l'invite à un repas | Avec mamie et papi |
| C'est une très belle nichée | Elle déguste un nougat | Les couleurs de ce sari | Mon cardigan en pilou |
| C'était un bon banquet | Je prendrai le pari | La hauteur de ces silos | J'ai adoré son récit |
| Ce n'est qu'un vain bouquet | On l'attache au piquet | Les liqueurs et les sodas | Elle a chassé les souris |
| C'est son meilleur côté | J'ai atteint le quota | Toutes les fleurs de la vallée | Le nombre deux au carré |
| Un authentique pâté | J'escalade le rocher | Ils refusent tous nos combats | Il ne met plus le gilet |
| Il est en grand danger | Il n'a pas de tabou | Cette vieille danse est une bourrée | Les pommes de terre en purée |
| Les plus anciennes momies | La leçon de tango | Les renards et les poulets | Il inhalait une bouffée |
| C'est le meilleur confit | Je t'ai pris ton valet | Les mitaines et les bonnets | Il a sauté le fossé |
| Il a un grand képi | Il adore le sofa | Faire des courses chez le boucher | Elle en a mis une pincée |
| C'est un gentil petit | Une odeur de forêt | Remplis donc un bon pichet | Elle débordait de bonté |
| C'est un énorme devis | Il enjambe le muret | Le mouton et le biquet | Elle a reçu un paquet |
| Elle a un plein caddie | L'érosion des galets | Elle a fait toute la montée | Ils ont payé leurs tickets |
| Elle fait un bon boulot | Ils écrivent un sonnet | Le dîner ou le souper | La frénésie du musée |
| Elle a pris 10 kilos | Le carrosse du cocher | L'explosion de la fusée | Ils ont péché des rougets |
| C'est le dernier fagot | C'est déjà du passé | L'arrivée de la rosée | Elle a commis un délit |
| C'est le plus beau mambo | Elle cultive sa beauté | Fromage blanc avec coulis | Il a mangé le kaki |
| Elle a de gros sabots | Ils dévorent leur goûter | La passoire et le tamis | Bien enrouler le tapis |
| Il ferme le grand verrou | J'ai reçu un colis | De la barde pour le rôti | Nous rentrerons le lundi |
| Le délicieux ragoût | Ils ont fait des semis | Le spectacle est le jeudi | Accroche-toi bien au garrot |
| C'est pas le bon menu | Ça sent bien le roussi | Le cheval part au galop | Il a grondé le minot |
| Le délicieux café | Travailler sans répit | Le fenouil et le pavot | Il nous faudrait un pivot |
| C'est un très bon début | Une question de culot | Change la lame du vieux rabot | Elle a ôté ses bijoux |
| Elle est en bonne santé | Le guidon du vélo | La souris fuit le matou | Il escalade le talus |
| Le dangereux combat | Le bienfait du repos | C'est une quête pour le salut | Tous les glaçons ont fondu |

Note. Les syntagmes sont tirés du matériel expérimental de Cason, Astésano, et al. (2015)

ANNEXES – PARTIE 2

voir le document complémentaire

Titre : 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

Mots clés : rythme, parole, sensorimotricité, supramodalité du rythme, amorçage, temps d'initiation de la parole

Résumé : Le rythme dans la parole est un phénomène complexe qui repose sur deux processus hiérarchiques pour décoder efficacement le signal continu de parole et permettre l'accès au sens : la segmentation en groupements pertinents (constituants prosodiques) et l'extraction d'une pulsation (réurrence des prééminences accentuelles). Les prééminences accentuelles attirent et guident l'attention de l'auditeur de syllabe accentuée en syllabe accentuée, instaurant une structure métrique sous-jacente qui permet d'anticiper les prééminences à venir, ce qui facilite leur traitement. Le paradigme de l'amorçage rythmique repose sur ce principe : les rythmes attentionnels de l'auditeur se couplent sur les périodicités d'un signal rythmique externe complexe (parole, musique), installant des repères temporels et métriques qui facilitent la perception et induisent également une structure métrique qui organise la production de parole subséquente.

Lorsque le rythme du flux de parole d'un locuteur ou les capacités de synchronisation rythmique d'un auditeur sont affectés, le couplage entre un rythme externe (parole) et les rythmes attentionnels peut être impacté, ce qui peut affecter la communication. Certains protocoles utilisent ainsi le rythme comme outil de rééducation de la parole et du langage. Cependant, l'effet différentiel des modalités rythmiques utilisées (auditif, moteur, verbal, musical) n'est généralement pas investigué, ce qui nous a amené à questionner l'effet de différents amorçages rythmiques sur le temps d'initiation de la parole. De plus, afin d'explorer plus avant le caractère supramodal du rythme, nous avons aussi questionné l'existence d'une période de référence supramodale et l'hypothèse d'un partage de mécanismes cognitifs pour le traitement du rythme et du langage.

Pour répondre à ces questionnements, nous avons élaboré un protocole expérimental constitué de deux activités réalisées par un échantillon de 40 adultes francophones neurotypiques. Dans une première activité visant à caractériser les capacités de traitement du rythme, nous avons recueilli des mesures relatives au Tempo Moteur Spontané (TMS) (période de référence et stabilité) et à la Synchronisation Sensori-Motrice (SMS) (stabilité et précision de la SMS). Dans une seconde activité (tâche de lecture à haute voix après différents types d'amorçages rythmiques, modulés par l'engagement moteur), nous avons analysé les erreurs de lecture et le Temps d'Initiation (TI) de la parole.

Nos résultats rapportent une fenêtre temporelle proche entre le TMS moyen global (712 ms) et le TI moyen de la parole (873 ms), sans que ces mesures soient corrélées. Concernant l'effet de l'amorçage rythmique sur la facilitation de la production de parole, nos résultats montrent un effet facilitant d'un amorçage auditif avec engagement moteur du participant par rapport à un amorçage auditif sans engagement moteur, ainsi qu'un effet facilitant d'un amorçage de type non langagier par rapport à un amorçage langagier (effet significatif) et un amorçage musical (tendance). Nos résultats montrent une différence de sensibilité à la congruence métrique de l'amorçage selon la structure métrique de l'amorce. Enfin, concernant l'hypothèse de mécanismes cognitifs partagés entre le traitement du rythme et du langage, nos résultats montrent que les participants les plus stables et précis dans les tâches de TMS et de SMS présentent un TI le plus lent et un nombre d'erreurs de lecture moindre (tendance).

Notre étude encourage ainsi à intégrer différentes modalités rythmiques comme perspective en remédiation de troubles de la parole, notamment l'amorçage rythmique auditif avec engagement moteur et l'amorçage rythmique de type vocal non langagier. Notre étude encourage d'autre part à affiner les hypothèses à propos de mécanismes cognitifs partagés entre rythme et langage, dans le but de mieux tenir compte des capacités de traitement du rythme pour l'évaluation des capacités de traitement du langage.

Title: From rhythm to speech: effect of linguistic, non-linguistic, and musical rhythmic primings modulated by motor engagement on speech initiation time

Key words: rhythm, speech, sensorimotor integration, rhythm supramodality, priming, speech initiation time

Abstract: As a complex rhythmic phenomenon, speech relies on two hierarchical processes to effectively decode the continuous speech signal and allow access to meaning: segmentation into relevant groupings (prosodic constituents) and extraction of a beat (recurrence of accentual prominences). Rhythm in speech thus provides valuable cues for segmenting the auditory stream and also structures production. Accentual prominences attract and guide the listener's attention from stressed syllable to stressed syllable, establishing an underlying metric structure from which the listener can anticipate the location of upcoming prominences, facilitating their processing. The rhythmic priming paradigm is based on this principle: a listener's attentional rhythms couple with the periodicities of a complex external rhythmic signal (speech, music), establishing temporal and metric markers that facilitate perception and also induce a metric structure organizing subsequent speech production.

In case of disturbances of a speaker's speech rhythm or a listener's rhythmic synchronization, the coupling between an external rhythm (speech) and attentional rhythms may be impacted, which may affect communication. Some protocols thus use rhythm as a tool for speech and language rehabilitation. However, the differential effect of the rhythmic modalities used (auditory, motor, verbal, musical) is generally not compared. In this study, we propose to investigate this question, to identify the most facilitating rhythmic priming for speech initiation. Moreover, to further explore the supramodal nature of rhythm and its impact on speech and language, we also questioned the existence of a supramodal referent period and the hypothesis of shared underlying cognitive mechanisms for rhythm and language processing.

These questions were tested with a sample of 40 French-speaking adults through two experimental activities. In Activity 1 (characterization of rhythm processing abilities), we collected measures related to Spontaneous Motor Tempo (SMT) (referent period and stability) and Sensorimotor Synchronization (SMS) (SMS stability and precision). In Activity 2 (reading aloud after rhythmic priming), we observed the number of errors and Speech Initiation Time (SIT).

Our results report a close temporal window between the average SMT (global average SMT = 712 ms) and SIT (global average SIT = 873 ms), albeit not correlated. Regarding the effect of rhythmic priming on facilitating speech production, our results show a facilitating effect of auditory-motor priming compared to auditory priming alone, as well as a more facilitating effect of non-linguistic priming compared to linguistic priming (significant effect) and musical priming (trend). Our results show a difference in sensitivity to the metric congruence of the priming, depending on the metric structure of the prime. Finally, regarding the hypothesis of shared cognitive mechanisms between rhythm and language processing, our results show that the participants who are the most stable and precise in SMT and SMS tasks show the slowest SIT and make the fewest reading errors.

Thus, our study encourages the integration of different rhythmic modalities in speech disorder remediation perspectives, particularly auditory rhythmic priming with motor rhythms and non-linguistic vocal rhythmic priming. Furthermore, our study encourages continued research on the hypothesis of shared cognitive mechanisms between rhythm and language processing to better integrate measures of rhythm processing abilities in language assessment.