UNTYING THE KNOT OF STUTTERING

Investigating cognitive profiles of stutterers through case-studies

---

Amélie VAN THORRE

Thèse présentée en vue de
l'obtention du grade de Docteur
en sciences psychologiques et
de l'éducation

Promoteur
Pr. Gilles VANNUSCORPS (UCLouvain)
Pr. Marie-Pascale NOËL (UCLouvain)

Président du jury

Pr. Alexandre HEEREN (UCLouvain)

Membres du jury

Pr. Arnaud Szmalec (UCLouvain)

Pr. Ivana DIDIRKOVA (Université Paul-Valéry Montpellier 3, France)

Pr. Anne-Lise LECLERCQ (ULiège)

Collection de thèses de l'Université catholique de Louvain, 2025

To Sélène et Lyra,

Table des matières

Table des matières ix

Résumé xv

Abstract xvii

Remerciements xix

General introduction 1

Chapter 1. Overview of stuttering 3

1. Causality and stuttering *6*

2. Speech production models as a framework for understanding stuttering *9*

2.1 Levelt's (psycholinguistic) model of speech production 10

2.2 (GO)DIVA (sensorimotor-based) model of speech motor control and development 12

2.2.1 Development of speech in infancy 13

2.2.2 Planning and execution in mature speech 14

3. Major cognitive theories of stuttering *18*

3.1 Evidence supporting the major theories of suttering 18

3.1.1 Stuttering as psycholinguistic disorder 18

3.1.2 Stuttering as an auditory/audio-motor integration disorder 20

3.1.3 Stuttering as a somatosensory/somatosensory-motor integration disorder 22

3.1.4 Stuttering as a rhythmic/timing disorder 24

3.1.5 Stuttering as a motor disorder 25

3.1.6 Stuttering as a multi-sensory-motor integration disorder 26

3.2 Challenges facing the major theories of stuttering 27

4. Stuttering or stutterings? *31*

4.1 Examining evidence of cognitive heterogeneity 32

4.1.1 Computational Modeling 32
4.1.2 Responses to treatments 33
4.1.3 Performance variability on cognitive processes 34
4.1.4 Neurobiology 34
4.2 Classifying stuttering into subtypes 38
5. Rethinking stuttering research *40*
5.1 Limitations of group study designs in advancing stuttering research 40
5.2 Advancing stuttering research through single-case methodology 43
5.2.1 How the single-case approach can improve our understanding of the cognitive basis of disorders 43
5.2.2 Benefits of the single-case approach in acquired disorders 47
5.2.3 Benefits of the single-case approach in developmental disorders 48
5.2.4 Implementation of the single-case approach to stuttering 49
6. Research question *50*
7. Outline of the thesis *50*
Chapter 2. Bégaiement et proprioception **53**
1. Introduction *54*
2. La proprioception dans le bégaiement *55*
2.1 Les données comportementales 55
2.1.1 Les études portant sur la reconnaissance de formes 55
2.1.2 Les études portant sur l'acuité proprioceptive 56
2.1.3 Les études portant sur la perturbation du feedback proprioceptif 57
2.2 Les données neurofonctionelles 58
3. Discussion *59*
Chapter 3. Is stuttering confined to articulated speech? **63**
1. Introduction *64*
2. Methods *67*
2.1 Participants 67

2.2 General procedure 70
2.2.1 Evaluation of stuttering severity 71
2.2.2 Stimuli 72
2.2.3 Tasks 73
2.2.4 Procedure 74
2.2.5 Scoring 75
3. Analysis and results *75*
4. Discussion *83*
5. Conclusion *84*
Chapter 4. Testing theories of stuttering: cognitive heterogeneity but shared orosensory tactile weaknesses **87**
1. Introduction *88*
1.1 Psycholinguistic abilities 90
1.2 Auditory skills 91
1.3 Somatosensory skills 92
1.4 Rhythmic and timing skills 93
1.5 Motor execution 94
2. Methods *95*
2.1 Participants 95
2.2 General procedure 98
2.3 Tasks and measures 100
2.3.1 Evaluation of stuttering severity 100
2.3.2 Linguistic assessment: efficiency of access to successive levels of representation 101
2.3.3 Linguistic assessment: inhibitory control 104
2.3.4 Auditory abilities 105
2.3.5 Somatosensory abilities 108
2.3.6 Rhythmic/timing abilities 114

2.3.7 Motor abilities 116

3. Analyses and results *117*

3.1 General analyses 118

3.1.1 Logic guiding the analysis and interpretation of the data 118

3.1.2 Criteria for the definition of weaknesses, strengths and dissociation 119

3.2 Group level analyses 120

3.3 Outlier detection and exclusion 120

3.4 Missing data 121

4. Results *121*

4.1 Linguistic assessments: efficiency of access to successive levels of representation 121

4.2 Linguistic assessment: inhibitory control 124

4.3 Auditory abilities 125

4.4 Somatosensory skills 128

4.5 Rhythmic/timing skills 134

4.6 Motor abilities 137

4.7 Summary of the results 140

5. Discussion *142*

5.1 Cognitive heterogeneity 142

5.2 A possible contribution of a tactile orosensory disorder? 148

6. Conclusions *155*

7. Acknowledgement. *158*

8. Supplementary materials *159*

General discussion **161**

1. Summary of the results *162*

2. Cognitive heterogeneity and its implications for the search of causation *164*

2.1 The importance of setting a logical framework 167

2.2 The causal implication of cognitive heterogeneity 168

3. Orosensory tactile weaknesses and stuttering: a cause or a consequence? *173*

3.1 Orosensory tactile weaknesses as a potential cause of stuttering 173

3.1.1 The potential role of tactile difficulties in the emergence of stuttering 174

3.1.2 The potential role of tactile difficulties in the disruption of the timing network that supports fluent speech 177

3.2 Orosensory tactile weaknesses as a potential consequence of stuttering 179

4. Limits and future perspectives *180*

5. Conclusion *183*

References **185**

Appendix **237**

1. Supplementary Information *237*

2. Performance consistency *255*

Résumé

Parmi les processus cognitifs nécessaires à une parole fluente, lesquels dysfonctionnent et sont à l'origine du bégaiement ? Cette question essentielle n'a toujours pas trouvé de réponse claire, et ce, malgré plusieurs décennies de recherche. À l'heure actuelle, la plupart des chercheurs s'accordent sur le fait que les personnes qui bégaient (PQB) présentent des profils cognitifs hétérogènes se marquant par des combinaisons variées de forces et de faiblesses au sein des habiletés qui sous-tendent la parole fluente. Cependant, et en dépit de ce consensus, cette reconnaissance a eu, jusqu'à présent, une influence limitée sur la manière dont le bégaiement est étudié. Cette thèse vise à combler cette lacune en apportant, pour la première fois, une démonstration empirique de l'existence de profils cognitifs hétérogènes chez les adultes qui bégaient (AQB). Pour ce faire, deux études complémentaires, reposant sur une approche d'étude de cas multiples, ont été menées. La première étude vise à déterminer si les PQB présentent des difficultés spécifiques au niveau de la mémoire de travail phonologique, c'est-à-dire au sein de processus indépendants de la production articulatoire de la parole ; tandis que la seconde évalue les performances individuelles de douze AQB dans cinq domaines essentiels à la parole fluente, à savoir les habiletés linguistiques, auditives, somatosensorielles, rythmiques et motrices.

Les résultats révèlent des profils cognitifs distincts, chaque participant présentant une combinaison variée de forces et de faiblesses au sein des domaines testés. Cependant, et de manière surprenante, un trouble de la perception tactile orofaciale est observé chez chaque individu. Cette difficulté orosensorielle pourrait ainsi constituer à la fois une condition nécessaire et suffisante à l'émergence du bégaiement.

Ces résultats ouvrent des perspectives majeures pour les recherches futures sur la compréhension des mécanismes cognitifs sous-tendant le bégaiement. Premièrement, ils indiquent que la perception tactile, largement négligée jusqu'ici, mérite une attention accrue, notamment pour déterminer si les difficultés observées relèvent d'un facteur causal ou d'une conséquence du trouble. Deuxièmement, l'hétérogénéité cognitive identifiée chez les AQB souligne l'importance d'évaluations individuelles plus complètes, afin de mieux prendre en compte la diversité des profils, tant dans l'analyse des facteurs impliqués dans le bégaiement

que dans la conception des interventions. Dans l'ensemble, ces conclusions invitent à repenser les approches théoriques, méthodologiques et cliniques du bégaiement, afin d'y intégrer pleinement l'hétérogénéité cognitive de la population des PQB.

Abstract

Which cognitive processes essential to fluent speech are disrupted and give rise to stuttering? After decades of investigation, we have yet to provide a clear answer to this critical question. Researchers widely recognize that people who stutter (PWS) exhibit heterogeneous cognitive profiles, characterized by distinct patterns of cognitive strengths and weaknesses across the broad range of abilities required for fluent speech. However, despite this consensus, this recognition has had little influence on how stuttering has been empirically investigated. The present thesis seeks to address this gap by providing the first empirical demonstration of cognitive heterogeneity among adults who stutter (AWS). Employing a multiple single-case design, two complementary studies were conducted. The first study aims to determine whether PWS exhibit specific weaknesses in phonological working memory, that is, within processes independent of speech production ; while the second assessed the individual performance of twelve AWS across five domains critical for fluent speech: linguistic, auditory, somatosensory, rhythmic, and motor abilities.

Our result reveals distinctive patterns of strengths and weaknesses across AWS in the domains tested. However, and surprisingly, weaknesses in orosensory perception emerge as a common feature, suggesting that these orosensory weaknesses may serve both as a necessary and sufficient condition for stuttering to emerge.

These findings raise important implications for future research that aims to elucidate the cognitive mechanisms underlying stuttering. First, they suggest that orosensory perception has received insufficient attention in previous investigations and warrants further study, notably, it remains to be determined whether observed orosensory weaknesses are a causal factor in stuttering or rather a consequence of the disorder. Second, the heterogeneous patterns of strengths and weaknesses observed among AWS highlights the need for more comprehensive individual assessment and careful consideration of diverse profiles in both etiological research and intervention development. Taken together, these results indicate the necessity for a broader perspective in stuttering research that integrates comprehensive cognitive assessment into theoretical models, research designs, and clinical practices.

Remerciements

En rédigeant cette dernière section de ma thèse, une série de flashbacks me traverse et m'offre l'occasion de retracer les événements marquants qui ont jalonné ces huit dernières années. Je repense d'abord à la personne que j'étais il y a huit ans, lorsque j'ai débuté cette thèse. Certainement très naïve et insouciante, je ne savais pas vraiment dans quoi je m'embarquais ; j'avais simplement l'envie d'explorer plus en profondeur certains aspects liés à la production de la parole et de contribuer à la compréhension et à la prise en charge de ces troubles. Être confrontée à la thématique du bégaiement s'est rapidement révélé challengeant. Je me suis très vite sentie noyée dans l'immensité que représente ce champ. Au fil du temps, j'ai souvent douté de ma capacité à mener ce projet à bien, tant les défis personnels et professionnels étaient importants. Historiquement, j'aurais eu tendance à abandonner, et j'y ai souvent songé ; je ne suis en effet pas une adepte des courses d'endurance, et je préfère parfois me préserver mentalement plutôt que persévérer face à l'adversité. C'est donc avec une certaine dose de joie et de fierté qu'en ce 17 octobre 2025, je clôture la rédaction de cette thèse. Si cela a été possible, c'est avant tout grâce à un nombre considérable de personnes. Je souhaite ainsi vous dédier ces quelques pages afin de vous exprimer toute ma reconnaissance pour tout ce que vous m'avez apporté et qui m'a permis de trouver la force d'arriver au bout de cette thèse.

Premièrement, je souhaite remercier mon promoteur, **Gilles Vannuscorps**. Merci d'avoir pris le risque de t'embarquer dans cette aventure de longue haleine avec moi, alors que nous n'avions eu l'occasion d'échanger que quelques fois via Skype (à l'époque). Tu as été un promoteur très présent, que ce soit lors de la première année, où nous n'étions que tous les deux dans le labo, ou lors de cette huitième année (et demi), où le labo a bien grandi et où tes domaines de recherche se sont également élargis. À chaque personne motivée tu offres ton temps, ton énergie mais aussi ta finesse d'analyse. Cet investissement de ta part a été précieux tout au long de ces années. Ton enthousiasme pour la recherche, ta rigueur ainsi que ton dévouement pour ton travail ont été à la fois source d'inspiration, de motivation, mais aussi parfois de doute. Eh oui, nos rendez-vous pouvaient se solder par un grand sentiment d'euphorie (même si je ne suis pas la personne la plus expressive), mais aussi par beaucoup

de remise en question et, parfois, le sentiment de ne pas être à la hauteur. Même si ces moments n'étaient pas toujours évidents pour moi, ils reflètent avant tout des fragilités qui me sont propres. Tu as toujours eu pour but de m'aider à m'améliorer et à évoluer. Je voulais également te remercier chaleureusement pour toute la bienveillance dont tu as fait preuve sur le plan personnel au cours de ces quatre dernières années. Conjuguer la réalisation de cette thèse et mon rôle de maman n'a pas toujours été facile. Tu as toujours été à l'écoute et su percevoir lorsque je me sentais dépassée ou que je n'arrivais pas à être pleinement disponible cognitivement. Tu m'as laissé le temps de rebondir à de nombreuses reprises et tu as toujours tenté de trouver des solutions pour me soulager quand cela était possible. Sans cela, il ne m'aurait réellement pas été possible de finaliser ce beau projet.

Je remercie également chaleureusement **Jean-François Patri**. Tu es arrivé en plein milieu du projet et l'investissement dont tu as fait preuve a été, pour moi, une réelle aide dans une période où je doutais de pouvoir mener à bien ce projet. Merci d'avoir pris le relais dans de nombreux domaines et de l'avoir fait avec autant de patience et de bienveillance.

Je suis également très reconnaissante envers **Marie-Pascale Noël**, **Arnaud Szmalec** et **Alexandre Heeren**, membres du comité d'accompagnement de cette thèse, qui, de par leur expertise respective, m'ont permis de prendre du recul mais aussi d'envisager mon travail sous un angle différent, enrichissant ainsi ma réflexion. Je vous remercie également pour vos encouragements et vos commentaires positifs tout au long de ces années. Je tiens aussi à vous remercier d'avoir accepté, avec **Anne-Lise Leclercq** et **Ivana Didirkova**, de faire partie de mon jury de thèse. Merci pour le temps et l'investissement consacrés à la lecture de cette thèse, ainsi que pour nos échanges durant la défense privée. Je me sens chanceuse d'avoir pu bénéficier de votre expertise, toujours dans un climat chaleureux.

La réalisation de cette thèse n'aurait pas été possible sans l'investissement et l'implication des **vingt personnes qui bégaient**, ainsi que des **quarante participant·e·s contrôles** qui ont accepté de me suivre dans cette aventure. Je vous remercie pour la confiance que vous m'avez accordée et pour votre disponibilité tout au long de ce projet. J'ai énormément apprécié parcourir ce bout de chemin à vos côtés.

Tout projet de recherche comporte son lot de défis. Le recrutement de participants a été, dans le cadre de cette thèse, particulièrement challengeant. C'est pourquoi je tiens à

remercier du fond du cœur **Julie Viardot**, **Cécilia Caravaca** et **Alice Hoffait** pour leur aide inestimable dans le recrutement et la collecte des données. Je remercie également les étudiantes qui ont participé, par le biais de leur mémoire de fin d'études, à la conception des tâches, ainsi qu'**Élise Rombaux** pour son implication dans la réalisation de la tâche linguistique.

Je souhaite également remercier mes nombreux collègues qui ont été, de près ou d'un peu plus loin, un réel soutien.

Plus particulièrement, je tiens à remercier mes compatriotes du Cogneurolab : **Antoine**, **Élise M.**, **Emmanuelle**, **Marie H.**, **Sarah**, **Suzanne** et **Virginie**. Vous avez toujours su trouver les mots justes pour proposer des pistes pertinentes permettant de faire avancer ce projet lors de nos lab meetings. Votre bonne humeur et votre positivité ont toujours été des sources de motivation. Sarah, je voudrais particulièrement te remercier d'avoir pris de ton temps pour m'aider dans mes analyses statistiques de dernière minute.

Je suis également très reconnaissante d'avoir pu parcourir ces huit dernières années en partageant mon temps de travail en tant qu'assistante. Sans cela, je n'aurais pas eu l'occasion de rencontrer les belles personnes qui composent ou ont composé l'équipe des assistantes logo : **Anne-Sophie**, **Auriane**, **Clémence**, **Émilie C.**, **Émilie J.**, **Élise B.**, **Élise M.**, **Élise V.**, **Joanne**, **Léa**, **Manon**, **Marie D.**, **Marion** et **Mathilde**. Merci à vous toutes d'avoir toujours été présentes pour prêter une oreille attentive dans les moments de joie mais aussi de doute ; de m'avoir encouragée et aidée à relativiser lorsque cela était nécessaire, que ce soit sur des aspects liés au travail ou sur des points plus personnels. Cela a été un réel plaisir de travailler à vos côtés. **Marie D.**, tu es la personne avec laquelle j'ai le plus collaboré durant ces huit dernières années. Je voulais te remercier plus particulièrement d'être la personne que tu es. Ça a été un plaisir de travailler avec toi et de partager les étapes de la maternité et de la vie en général. Tu as toujours été présente et un soutien infaillible, toujours sans jugement.

Je tiens également à remercier **Claire**. Ton attention envers les autres est précieuse et je te remercie sincèrement pour ton soutien, tout spécialement durant ces derniers mois qui ont été particulièrement intenses.

L'aboutissement de cette thèse n'aurait certainement pas été possible sans l'amitié précieuse de **Sophie B**. Qui aurait pu prédire que partager notre bureau durant nos deux premières semaines au sein du troisième étage allait mener au développement d'une si belle amitié ?

Sophie, tu fais partie de ces personnes uniques à qui l'on peut tout confier sans se sentir jugée et avec qui on peut être totalement authentique. Tu sais toujours trouver les mots justes pour me faire voir les choses telles qu'elles sont et ainsi me permettre de me remettre en question ou de relativiser. Tu as été une présence continue, tant sur le plan professionnel que personnel. Nous avons vécu ensemble des étapes importantes de la vie, et les traverser avec toi a permis de rendre les choses plus douces. Je te remercie aussi chaleureusement pour avoir joliment illustré la couverture de cette thèse !

Je suis également profondément reconnaissante envers **Marie-Pierre**. Merci d'avoir été une professeure si inspirante durant mes années d'études, de m'avoir accompagnée et guidée pendant mon stage, mais aussi de m'avoir donné l'opportunité de me lancer dans cette grande aventure qu'est la thèse. J'ai énormément apprécié nos riches discussions et le fait de travailler à tes côtés.

Il me reste à remercier l'ensemble de mes compatriotes du troisième étage et représentants assistants : **Nicolas**, **Laurie**, **Michaël**, **Benjamin**, **Samuel**, **Sylvain**, **Laure**, **Lucas**, **Solène**, **Virginie**, **Jolijn**, **Marie V.**, **Chloé**, **Hubert**, **Elisa**, **Déborah**, **Catherine** et **Céline**. Merci pour les moments conviviaux, les discussions rapides (ou non) dans les couloirs, ainsi que pour vos encouragements tout au long de ces années. Je remercie également chaleureusement l'ensemble de l'équipe administrative de la faculté pour leur disponibilité et leur efficacité à répondre à la moindre de nos questions, toujours avec le sourire, merci **Anne, Dominique, Doriane, Martine, Nadine, Nathalie, Régeane, Roland, et Siegrid**.

En plus de cet entourage professionnel, j'ai également pu compter sur mon entourage amical et familial. Je remercie particulièrement **Alicia, Amélie, Armelle, Estelle, Florence, Tracy, et Valentine**. Malgré les années et la distance qui ont pu s'installer à cause de nos vies professionnelles et/ou la maternité, vous êtes toujours présentes pour partager quelques mots et des moments conviviaux, que ce soit dans les moments de joie ou lors des périodes plus difficiles. Vous faites vraiment partie des personnes les plus importantes dans ma vie et je me sens extrêmement chanceuse d'avoir croisé votre route à un moment donné. Je remercie également mes amis rencontrés, soit au détour de moments guindaillesques, soit autour d'une bonne tasse de thé, ou encore au gré de la vie : **Amandine S., Alexis, Amaury, Aurélie, Bruno, Caroline, Christophe, Jean, Jean-Benoît, Line, Maely, Marie-Laure, Marine, les deux Martin D., Noémie, Olivier, Paul, Quentin, Sophie H.,** et **Thibaut**. Même si nos emplois du temps (et l'âge avançant ?) ne nous permettent plus de faire la fête jusqu'au bout

de la nuit ou de partager des orgies de thé aussi souvent qu'auparavant, c'est toujours un plaisir de vous retrouver et de passer d'agréables moments ensemble. Au-delà de votre amitié, un grand nombre d'entre vous n'a pas hésité à se porter volontaire pour permettre que cette thèse voie le jour. Passer ces dix heures de testing au côté de chacun d'entre vous a rendu cette expérience plus douce et je vous suis sincèrement reconnaissante.

Il me tient maintenant à cœur de remercier profondément ma famille : merci **Papa**, **Maman** et **Guillaume** pour votre présence, votre confiance et votre soutien depuis 34 ans déjà. Je sais que, quoiqu'il arrive, je peux compter sur vous et que vous serez toujours présents pour moi, que ce soit dans les bons ou les moins bons moments. Ces derniers mois, vous m'avez soutenue et encouragée à me dépasser, et je suis à présent fière de pouvoir partager avec vous la joie de clôturer ce chapitre de ma vie et d'ouvrir le suivant à vos côtés. Je remercie également mes **grands-parents** pour leur soutien. Je ne pourrai malheureusement pas fêter ce moment avec trois d'entre vous, mais j'ai toujours une petite pensée pour vous où que vous soyez.

Trois personnes sont encore à remercier, et non des moindres ! Il y a quatre et deux ans, j'ai eu la chance de devenir la maman de deux adorables petites filles. **Sélène** et **Lyra**, vous avez bouleversé ma vie à bien des niveaux. Non seulement je me sens grandie et plus épanouie avec moi-même depuis votre arrivée, mais je considère également la vie différemment. Grâce à vous, je me sens davantage connectée à ce qui importe réellement dans ce monde, ce qui a parfois rendu la poursuite de cette thèse difficile, notamment lorsque je peinais à retrouver du sens à mon travail, surtout après mes retours d'arrêt de maternité. Alors oui, vous ne pourrez pas lire ces quelques mots avant plusieurs années, néanmoins je voulais vous remercier, tout simplement, d'être vous et de combler ma vie (et mes nuits !) par votre présence. Je me sens chanceuse et très fière d'être votre maman. Ces derniers mois, il n'a pas été facile de conjuguer ma vie de maman avec la finalisation de cette thèse. La fatigue m'a souvent fait perdre patience plus rapidement et j'ai été moins présente à vos côtés. Malgré votre jeune âge, vous avez su accepter ce changement dans notre quotidien. Il me tenait aussi à cœur de mener cette thèse à son terme pour vous, afin de vous montrer que tout est possible dans la vie et que, même face à une montagne qui paraît infranchissable, entourée des bonnes personnes, on peut y arriver. Et sachez que je serai toujours présente pour vous

aider à franchir toutes les montagnes qui se dresseront sur votre chemin. The last but not the least ! **Sébastien**, ces quatorze dernières années ont été chamboulées à bien des niveaux. On peut dire qu'après deux enfants et deux thèses, nous sommes prêts à relever tous les défis qui s'offriront à nous ! Tu as toujours trouvé le moyen de m'encourager à persévérer et, à bien des reprises, tu as même pris de ton temps pour mettre à profit ton expertise dans de nombreux domaines afin de m'aider (comme passer une heure à te faire lire et relire la première phrase de l'abstract de cette thèse parce que je ne parvenais pas à trouver les mots justes). L'aboutissement de cette thèse n'aurait donc pas été possible sans toi et sans la confiance que tu m'as témoignée. Ces quatre dernières années, et plus particulièrement cette dernière année, ont été particulièrement intenses. Merci de m'avoir soutenue en mettant en place tous les aménagements possibles pour te dégager du temps et me libérer le mien. Je voulais tout simplement te remercier d'être toi et de partager avec moi tous les challenges et les moments de bonheur que la vie nous offre, j'ai hâte de voir ce que l'avenir nous réserve.

General introduction

The production of fluent speech is among the most complex motor skills humans must learn. While not a perfect analogy, since speech development is more intricate and biologically predisposed, speech learning shares some similarities with the process of learning how to drive a car. At first, both skills require intense focus and deliberate effort. When learning to drive, you must consciously think about each action: pressing the pedals, shifting gears, steering, checking mirrors, and following traffic rules. Similarly, when children learn to speak, they must actively work to integrate multiple sources of information, including linguistic, auditory, somatosensory, and motor representations. Additionally, producing complex patterns of speech sounds requires precise activation, timing, and coordination of numerous muscles in the respiratory, laryngeal, and articulatory systems, all while simultaneously processing the meaning of words.

However, with practice and repetition, these complex actions become automatic. Just as experienced drivers can navigate busy streets while carrying on a conversation, fluent speakers can produce intricate sentences without consciously thinking about each word or muscle movement. In both cases, what begins as a complex task eventually becomes second nature. Yet, for some individuals speaking represents a real challenge; for instance, individuals who stutter.

Stuttering is one of the oldest recorded speech disorders, with descriptions dating back to ancient Egyptian civilization around 2000 BCE (Brosch & Pirsig, 2001). Furthermore, stuttering is a relatively common condition, affecting approximately 1% of the population (Gattie, Lieven, & Kluk, 2025), and is found across all civilizations, regardless of language or cultural differences (Guitar, 2023). Given its longstanding recognition and widespread prevalence, stuttering has become one of the most extensively studied speech disorders. However, despite the significant contributions of numerous researchers and the amount of data collected over the past centuries, our understanding of the cognitive mechanisms underlying stuttering still remains limited. Although numerous theories have been proposed to explain the origins of stuttering, no single theory sufficiently accounts for the full range of difficulties observed in the disorder, indicating that stuttering likely arises from multiple factors.

In this context, and building on this body of work, the present thesis aims to investigate a broad range of components involved in fluent speech to determine whether they could contribute to stuttering. Furthermore, it examines the extent to which these factors may vary across people who stutter (PWS). In this context, Chapter 1 presents the foundational context that led to our research question. The core aim of this chapter is to explore why, despite decades of extensive research, our understanding of the cognitive causes of stuttering remains limited. To address this, I will first propose a logical framework for conceptualizing the mechanisms underlying developmental disorders such as stuttering. Next, I will examine the mechanisms involved in fluent speech that have been identified as potential contributing factors to stuttering, analyzing these in light of the major theoretical hypotheses advanced in the literature. I will then review evidence suggesting that stuttering is not a uniform condition, but rather is composed of cognitively heterogeneous individuals. Finally, I will propose and discuss the utility of an alternative framework for studying stuttering, the multiple single case approach.

Chapter 1. Overview of stuttering

Given the prevalence of stuttering, it would be nearly impossible to encounter someone who has never heard of this disorder. In fact, when I began my Ph.D. and explained to those around me that my research would focus on understanding the mechanisms underlying stuttering, many looked puzzled and would respond with remarks such as: “Oh, that’s a topic I can understand. But don’t we already know that, isn’t it caused by…?”.

And indeed, it is true that we have gathered a substantial body of knowledge on stuttering. However, we are still unable to piece it together into a coherent and comprehensive story. A simple example of the mystery surrounding stuttering is the lack of an unanimously accepted definition (Alfonso, 1991; Bloodstein, Ratner, & Brundage, 2021; Perkins, 1990; Wingate, 1964). While this may seem trivial, the absence of a clear definition of stuttering has significant implications; not only in terms of the misconceptions that persist about the disorder but also for research and clinical practice (Yairi & Seery, 2021).

If you are looking for information about the disorder online, you might come across descriptions stating that stuttering is a communication disorder: “people generally do not stutter when speaking alone” (APB, n.d.), or “people stutter particularly in the presence of an interlocutor” (Thérapeutes, 2025). Does it mean that stuttering is restricted to social anxiety? You might also read that stuttering is a “genetic disorder” (BBC, 2020), or a “neurological disorder” (BBC, 2020), or an “emotional disorder” (Jones et al., 2014). However, it does not mean that these predispositions (genetic, neurological or emotional) are sufficient for stuttering to occur (Packman & Attanasio, 2017). Additional conditions, such as impairments in the processes required for fluent speech, might also be required.

Published definitions themselves are often divergent. Some describe stuttering as a "fluency disorder" (American Psychiatric Association Dsm-5 Task Force, 2013) without specifying what constitutes typical fluency. Others mention "abnormal tension" of the vocal, oral or respiratory musculature (Starkweather, 1995; Tichenor et al., 2018). Alm and colleagues (2025) propose that such abnormal tension arises as a consequence of motor blocks, specifically, positive motor signs, reflecting a failure to transition smoothly between speech movements. During such motor blocks, PWS may engage their musculature more

forcefully in an effort to break through a fixed articulatory posture and resume speaking. This sustained, moderate muscular tension represents an ongoing attempt to overcome the block. If tension continues to rise, these contractions can escalate to involuntary tremors (rapid rhythmic contractions) or other physical concomitants. However, does this mean that every stuttering behavior should be associated with muscular tension in order to be determined as a stuttering event? Terms like "effortful speech" (Ingham et al, 2006) are also used, yet lack clear operational definitions. Additionally, some definitions propose possible etiologies, such as a “speech motor control disorder” (Civier, 2010) or “motor timing disorder” (Alm, 2004). Does this imply that these impairments must be corrected? If so, what tools should be used?

Consequently, most published definitions err on the side of caution and adopt a behavioral perspective, describing the features or symptoms of stuttering primarily from the listener’s point of view. These definitions can generally be summarized as follows: *stuttering is commonly characterized by interruptions in the forward flow of speech due to disruptions in the utterance of sounds, syllables, or words* (Alfonso, 1991; Alm, 2005; Perkins, 1990; Van Riper, 1982; Wingate, 1964), with some definitions also referring to phrase-level disruptions. However, there is no consensus on the specific manifestations of these disruptions (Yarus, 1997). They can take various forms, including blocks, prolongations, repetitions, abnormal muscular tension, and involuntary movements such as tremors, eye blinks, or tics (typically involving the mouth, throat, and face), as well as breathing incoordination (Alonso, 1991; Alm, 2005; Sheikhbahaei et al., 2023; Wingate, 1964). Furthermore, definitions often lack precision in describing these features of stuttering, making it difficult for researchers and clinicians to determine exactly what they should be assessing (Bloodstein et al., 2021). Some lists of possible stuttering behaviors are so extensive that distinguishing a PWS from individuals with other disorders becomes challenging. We joined Frith (2001, p.557) who state that “defining developmental disorders at the behavioral level is unsatisfactory as behavioral definitions are variable and subject to noise from many factors that are not germane to the question of developmental disorders”. In addition to these speech-related difficulties, numerous comorbid behaviors have been identified, such as (social) anxiety, avoidance, emotional disorders, ADHD, and obsessive-compulsive disorder (OCD) (Choo, Smith, & Li, 2020).

Beyond the diversity of these behaviors, stuttering is marked by considerable variability, both within and between individuals. Intra-individual variability refers to fluctuations in severity, with stuttering being absent in some situations and pronounced in others, or changing across periods of life (Tichenor & Yaruss, 2018). Inter-individual variability refers to the differences in how stuttering manifests from one person to another. Many stutterers exhibit only a subset of the wide range of symptoms typically associated with stuttering (Yairi, 2007). While some individuals present with many of these features, others may show only mild disfluencies without observable struggle or tension. For some, stuttering is accompanied by comorbid and secondary behaviors such as physical tension, anxiety; others may conceal their symptoms so effectively that avoidance behaviors and social anxiety become the most salient manifestations, a presentation known as "masked stuttering" (Sheikhbahaei et al., 2023). The difficulty in establishing a specific syndromic feature of stuttering has even led some authors to question whether what we call stuttering truly exists or if it merely represents "dysrhythmic speech symptoms" of other conditions (St. Onge & Calvert, 1964).

In this thesis, the operational definition of stuttering adopts the behavioral description noted above: *Stuttering is commonly characterized by interruptions in the forward flow of speech due to disruptions in the utterance of sounds, syllables, or words*. *These disruptions can take various forms, including blocks, prolongations, repetitions, abnormal muscular tension, and involuntary movements such as tremors, eye blinks, or tics (typically involving the mouth, throat, and face), as well as breathing incoordination (Alonso, 1991; Alm, 2005; Sheikhbahaei et al., 2023; Wingate, 1964)*. However, given the substantial variability both within and between individuals in the manifestation of stuttering, additional criteria were necessary for determining participant inclusion. Accordingly, individuals were classified as PWS if they met the following criteria: (1) they had received a formal diagnosis of stuttering from a speech-language therapist; (2) they reported experiencing stuttering-related difficulties in their daily life; and (3) they self-identified as a person who stutters at the time of testing. This approach was adopted because some PWS may not exhibit overt stuttering behaviors during assessment, either due to the effective use of speech therapy techniques, the context-dependent nature of stuttering, or other individual factors. By combining clinical diagnosis, self-report, and self-identification, we aimed to ensure the validity and relevance of participant selection, despite the behavioral variability inherent in stuttering.

What failures in the processes that govern fluent speech give rise to the broad manifestations observed in stuttering? To move toward a deeper understanding of the disorder, it is essential to examine its underlying causes. This requires shifting from a descriptive account of symptoms to a causal framework that can explain how and why stuttering emerges and persists in some individuals but not others. The following section explores how the concept of causation can be addressed in the field of stuttering research.

1. Causality and stuttering

When confronted with something they do not understand, human beings have an inherent need to find an explanation. Stuttering is no exception to this rule. Thus, from early discussions dating back to the 4th century BCE to the present day, identifying what causes stuttering has been a central focus for many philosophers and scientists.

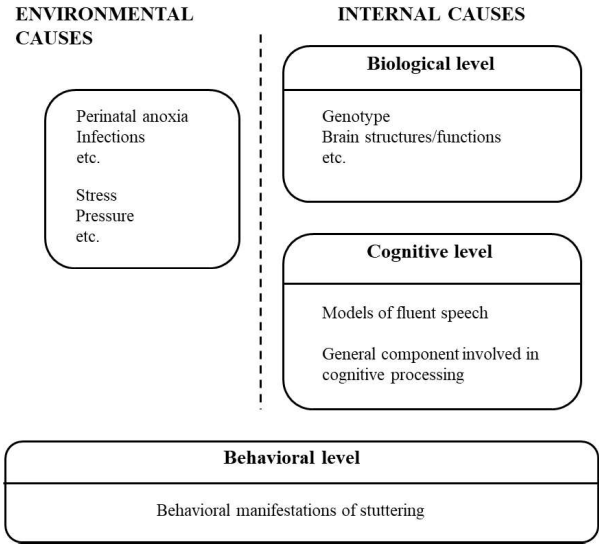
This section does not aim to be a philosophical essay on the nature of cause or causation. This question has been the focus of many philosophers over the centuries, making the philosophical literature on causation dense with debates and diverse perspectives (see White, 1990). This topic could, therefore, be the central focus of an entire thesis, which is not the case here. Rather, the goal is to establish a logical framework to support the understanding of the mechanisms underlying developmental disorders such as stuttering. We join Packman and Attanasio (2017) with the view that reasoning from a causal perspective helps generate new ideas, guide scientific research, and provides a framework for understanding relationships between events particularly when discussing causal knowledge of complex phenomena like stuttering. Failing to adopt a structured causal approach can lead to misguided reasoning, ultimately resulting in erroneous conclusions about the causal relationship between an event and stuttering (see Attanasio, Onslow, & Packman, 1998; Packman & Attanasio, 2017, for review). One example of misguided reasoning is the post hoc fallacy, in which one assumes that because one event follows another, the first event must have caused the second. Simply put: event A happened before event B; therefore, A caused B. For instance, PWS often report experiencing a traumatic event shortly before the onset of their stuttering and subsequently conclude that trauma caused the disorder. However, just because trauma precedes the onset of stuttering does not mean it is its cause. Several key observations challenge this assumption: (1) many PWS do not report a traumatic event prior

to the onset of their stuttering, indicating that trauma is not a NECESSARY CONDITION for stuttering to develop. (2) Many individuals who experience trauma do not develop stuttering, suggesting that trauma alone is not a SUFFICIENT CONDITION to cause stuttering.

Packman and Attanasio (2017) also discuss the idea that causality could be seen as a chain of events. In line with this perspective, Jackson and Coltheart (2013) have proposed a causal modeling framework, based on the approach of Morton and Frith (1995, see also Frith, 2001), to explore individual differences in developmental disorder (taking reading development as an example), a model that could be adapted for the study of stuttering. In their framework, three levels of analysis are proposed, which we have adapted to stuttering (see Figure 1): (1) a behavioral level, the observable manifestations of stuttering; (2) a cognitive level, explanatory hypotheses based on models of fluent speech that account for the behaviors observed in stuttering; (3) a biological level, the explanations based on neurobiological differences. Additionally, environmental influences or life experiences could impact the biological and cognitive levels.

Figure 1

Causal framework of stuttering



Note. Causal framework of stuttering based on Jackson and Coltheart (2013)

Jackson and Coltheart (2013) argued that an explanation at these three levels would allow for a full comprehensive understanding of the profile of any individual with a developmental disorder. However, in their view, achieving a complete explanation is not necessary. The primary focus is on the link between level 1 (behavioral manifestations) and level 2 (cognitive dysfunctions), specifically identifying which dysfunctions in the cognitive system give rise to the observable symptoms of the disorder. They refer to these as PROXIMAL CAUSES, the direct mechanisms underlying the disorder. According to their perspective, uncovering these PROXIMAL CAUSES, through the cognitive neuropsychological approach, enables a better understanding of the behavioral manifestations of developmental disorders, independently of biological factors (level 3) and environmental influences. These latter causes are what they define DISTAL CAUSES, factors that contribute to but do not directly produce behavioral symptoms (but see Thomas & Karmiloff-Smith, 2002). Furthermore, Frith (2001) noted that some researchers consider a biological-level diagnosis of developmental disorders preferable, as it would be perceived as more objective and therefore less prone to disagreement. However, she argued that this assumption is far from secure, given that findings from genetic and other biological research remain inconsistent. To date, there is no definitive biological marker that can independently verify a diagnosis. Interestingly, Attanasio et al. (1998) suggest that "the failure to uncover the causes of stuttering" (Attanosio et al., 1998, p.265) is linked to the fact that most theories have focused on DISTAL CAUSES, which has consequently hindered progress in understanding the nature and causes of stuttering. According to them, this focus has led to complex notions that are difficult to test and validate scientifically.

In the present thesis we adopt the causal modeling framework proposed by Jackson and Coltheart (2013), adapted for stuttering, and focus on uncovering the PROXIMAL CAUSES of stuttering independently of the DISTAL CAUSES. The methodological framework we use to achieve this is inspired by the cognitive neuropsychology approach (see section 5). Additionally, we specifically examine whether any of the potential PROXIMAL CAUSES can be considered NECESSARY (present in all) and/or SUFFICIENT (the only detected deficit in some) factors in explaining stuttering in these individuals.

Before delving into the core of this thesis, it is essential to first present the major PROXIMAL CAUSES that have been proposed in the literature, along with the models of fluent

speech on which they are grounded. While this section outlined broader perspectives on causation, the next two sections focus on specific cognitive mechanisms that may be directly involved in the manifestation of stuttering. To lay this foundation, we begin with an overview of two influential models of speech production, which provide the conceptual and structural basis for many of the cognitive theories discussed later in this thesis.

2. Speech production models as a framework for understanding stuttering

To fully understand the cognitive theories proposed to explain stuttering, it is essential to consider the theoretical roots and foundational assumptions on which they are based. These theories, and by extension, the cognitive battery developed for the study presented in Chapter 4, were not constructed in isolation or from intuition alone, but are grounded in well-established scientific models of fluent speech production. Specifically, many of the mechanisms hypothesized to be impaired in PWS are derived from broader models of how speech is planned, encoded, controlled, and executed in typical speakers. Importantly, the cognitive theories of stuttering can be broadly categorized into those that emphasize psycholinguistic processes, such as lexical access, phonological encoding, or internal monitoring, and those that focus on sensorimotor-based mechanisms, such as speech initiation, feedback processing and articulatory sequencing. In line with this distinction, this section presents two influential models of speech production that serve as the conceptual foundation for our approach: Levelt's model of speech production (Levelt, 1989), which reflects a psycholinguistic framework, and the Gradient Order Directions Into Velocities of Articulators (GODIVA) model of speech motor control, developed as an extension of the DIVA model (Directions Into Velocities of Articulators), which together reflect a sensorimotor-based perspective on speech planning and execution (Bohland, Bullock, & Guenther, 2010; Guenther, 2016). Levelt's model of speech production has served as a foundational framework for theories emphasizing difficulties in the planning stages of speech production, particularly conceptualization and formulation. In contrast, the GODIVA/DIVA framework is often associated with theories focusing on motor execution and sensorimotor integration, such as articulatory sequencing and feedback control. However, it is important to acknowledge that the two frameworks both share a hierarchical organization, competitive sequencing and an interface between linguistic and motor processes, although at different

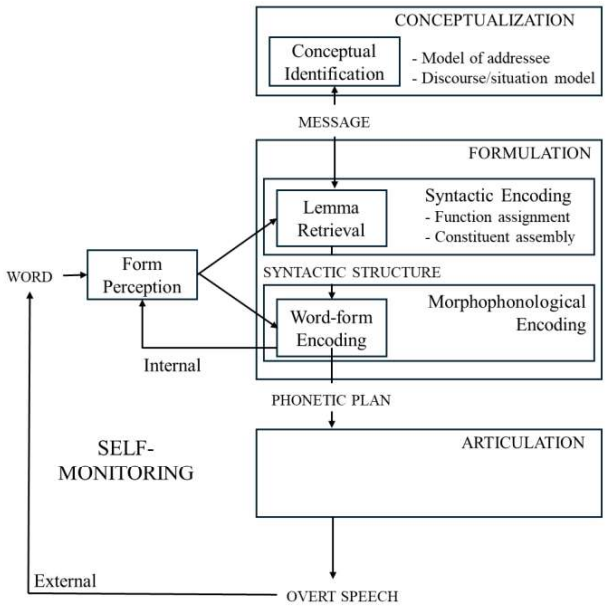
levels of specificity. While Levelt's model provides an abstract linguistic architecture, GODIVA/DIVA extends it with neurocomputational specificity, bridging phonological planning and sensorimotor execution. This integration underscores their complementary roles in understanding speech production.

2.1 Levelt's (psycholinguistic) model of speech production

Fluent speech production is achieved by progressing through multiple component processes and representations (Roelofs & Ferreira, 2019). Levelt (1989) and Levelt, Roelofs, and Meyer (1999) describe this process as being organized into three main processing stages (see Figure 2).

Figure 2

Levelt Model of speech production



Note. Cognitive architecture of speech production based on Levelt (1989, 1999) (Roelofs & Ferreira, 2019)

The first, CONCEPTUALIZATION, is a non-linguistic stage during which the speaker generates the message they intend to verbally express by selecting relevant concepts and organizing them into a coherent message-level representation. At this point, no words have been accessed; the output of this stage is purely conceptual and reflects only the intended meaning of the message to be formulated.

The second stage, FORMULATION, converts the conceptual message into a linguistic representation. This stage is lexically driven, meaning that the concepts selected during conceptualization activate corresponding words (lemmas) in memory, making available the syntactic and morphophonological information needed to construct a syntactic structure and phonetic plan (Roelofs & Ferreira, 2019). Thus, FORMULATION can be subdivided into three sub-processes: lexical selection, morpho-phonological encoding/syllabification and phonetic encoding. *Lexical selection* involves retrieving the appropriate lemmas from the mental lexicon along with their associated syntactic information. For example, in French, this includes grammatical gender, allowing the speaker to know whether a noun is masculine or feminine. According to Levelt (1989) and Levelt and collaborators (1999), lemma selection is governed by a statistical competition mechanism, in which the most highly activated lemma is selected. Once lemmas are selected, syntactic encoding allows for the assembly of syntactic constituents and the assignment of grammatical functions. *Morphophonological encoding and syllabification* involve retrieving the morphological and phonological information of the selected words, including the word's syllabic structure, what Levelt refers to as the "word-form" of a lemma (Levelt, 1989; Levelt et al., 1999; Roelofs & Ferreira, 2019). This stage converts linguistic units into their appropriate sound representations, ensuring that words are produced in the correct order, at the right tempo, and with appropriate prosody (Harley, 2013; Roelofs & Ferreira, 2019). Once the morphemes and phonemes are retrieved, a phonological word representation is formed and used to access corresponding syllables that are stored in a syllabary, a mental inventory of gestural patterns for the most frequently used syllables in a speaker's language (Levelt, 1989; Levelt et al., 1999; Roelofs & Ferreira, 2019). Once the syllables have been selected, the process moves to the final sub-stage of FORMULATION: *phonetic encoding*. Here, the corresponding syllabic motor programs are retrieved and assembled into a detailed articulatory plan that specifies the timing, coordination, and motor execution parameters required for speech (Roelofs & Ferreira, 2019).

The final main stage, ARTICULATION, carries out the phonetic plan by engaging the motor system to produce the articulatory movements that result in overt speech.

The model also specifies two self-monitoring loops that can detect errors during speech production (Levelt et al., 1999). The first is an external monitoring loop, which evaluates the speaker's overt speech output, during and after articulation, through auditory and somatosensory feedback and detects discrepancies between the intended message and the spoken output. The second is an internal monitoring loop, which inspects internal representations during the speech planning process. Based on the work of Wheeldon and Levelt (1995), internal monitoring is believed to occur as early as the phonological word representation stage, allowing for the detection and correction of errors before articulation begins. This internal mechanism enables speakers to edit their planned speech, such as cancelling, reformulating, or delaying speech onset, before execution.

Based on the description of the model (see Figure 2), the production of the word "cat" would proceed through as follow: The production of "cat" begins with the generation of a non-linguistic conceptual intention referring to a cat (CONCEPTUALIZATION). This activates lexical candidates (FORMULATION) such as "dog", "pet", and "cat", with "cat" ultimately selected based on its highest activation (*lexical selection*). The word-form <cat> is retrieved, and its phonemes /k/, /æ/, and /t/ are assembled and syllabified into a single CVC syllable (*morpho-phonological encoding*). The corresponding motor program is retrieved and structured into an articulatory plan (*phonetic encoding*), which is then executed to produce the word (ARTICULATION). Throughout, internal and external monitoring systems check for errors, allowing for correction either before or during speech production.

2.2 (GO)DIVA (sensorimotor-based) model of speech motor control and development

Since the early 1990s, Guenther and collaborators (Guenther, 1994; Guenther, 2016; Tournville & Guenther, 2011) have developed a neurocomputational model of speech motor production and control, known as DIVA. The model is capable of simulating both speech production and development by describing how speech sounds are transformed into articulatory movements. This model is biologically plausible, meaning it incorporates brain regions thought to be involved in generating the articulatory movements required for speech.

Over time, the model has been expanded to include a realistic articulatory synthesizer and a model of speech perception (Meier & Guenther, 2023). Among its many applications, the DIVA model is used to predict the effects of brain lesions by selectively disabling specific components of the model. In 2010, Bohland, Bullock, and Guenther expanded the DIVA model with the introduction of the GODIVA model, which adds a planning interface to account for how speakers plan, structure, select and gate the upcoming speech units of their utterances prior to articulation. Thus, fluent speech is achieved through a fine-tuned interaction between these two models that plan, select and execute what we want to say.

In this section, we first describe how infants develop speech and acquire mature speech production abilities based on the DIVA model. We then turn to a description of how mature speech production is carried out, drawing on an integrated GODIVA-DIVA model.

2.2.1 Development of speech in infancy

Speech production is no longer viewed as a purely motor act, but rather as the result of a dynamic interaction between the motor and sensory systems, specifically auditory and somatosensory (proprioception: sense of body position and tactile perception: the sense of touch) processing (Guenther & Hickok, 2015; Shamma et al., 2021). During the first two years of life, infants progress through several developmental stages (see Oller & Eilers, 1988) in order to acquire the mature sensorimotor representations needed for efficient and fluent speech (Guenther, 2016; Meier & Guenther, 2023). This process initially relies heavily on their ability to perceive speech sounds and their properties, such as frequency, intensity, duration, and acoustic features like formant structure. From the earliest stages, including in utero (Gervain, 2018), infants begin by processing the speech of their caregivers, which provides the initial auditory model for speech development. As they move through the early stages of phonation, such as the various stages of babbling, they increasingly begin to monitor their own vocalizations. Gradually, infants construct auditory-motor representations that correspond to the speech units of their native language. According to Guenther (2016; see also Meier & Guenther, 2023), speech acquisition initially relies on auditory targets: infants learn to match their vocalizations to the auditory patterns they hear. At a later stage, somatosensory targets, the tactile and proprioceptive sensations associated with speech movements, are acquired as infants gain more experience with producing speech sounds. These somatosensory representations are progressively integrated with auditory-motor

representations, resulting in a multisensory-motor representation for each speech sound. Unlike auditory representations, which can be formed through listening alone, somatosensory representations can only be developed through the infant's own speech production. Consequently, somatosensory targets are hypothesized to emerge later than auditory targets and are shaped by prior auditory experience. This developmental sequence implies a hierarchy among sensorimotor representations: auditory targets play a primary role in shaping early speech acquisition, while somatosensory targets are incorporated subsequently to refine and support accurate speech production. In summary, auditory experience guides the formation of somatosensory representations, leading to a layered, hierarchical organization of sensorimotor targets in speech development (but see Patri, Diard & Perrier, 2019, for a critic of hierarchical sensorimotor hypotheses).

Early speech development thus depends heavily on both auditory and somatosensory feedback mechanisms (Guenther, 2016). Importantly, somatosensory feedback plays a unique role in speech monitoring: because it provides internal information about articulator movement and contact, it can detect certain speech errors before articulation is completed, unlike auditory feedback, which can only respond to the acoustic signal after it has been produced (Guenther, 2016). Mature speech production is achieved once the individual has fully integrated motor programs for the speech units of their language, enabling direct access to these representations during speaking. At that point, auditory and somatosensory systems continue to serve a critical monitoring function, detecting errors and updating speech motor representations in response to anatomical changes over time (e.g., due to growth or aging). Meier and Guenther (2023), as well as Callan et al. (2000), provide detailed accounts of how the DIVA model simulates the learning process infants undergo to acquire speech.

2.2.2 Planning and execution in mature speech

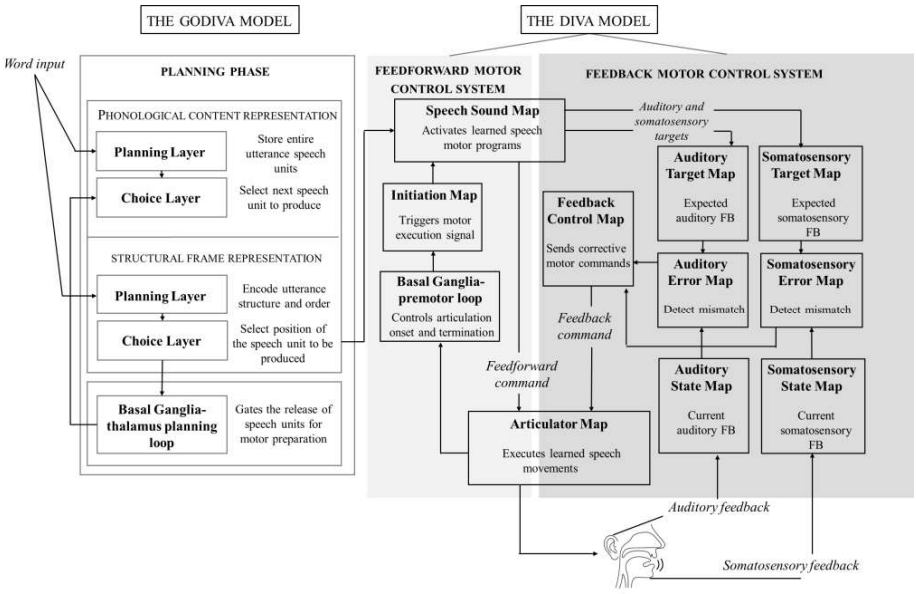
This subsection aims to provide a functional overview of mature speech production by integrating key components of the GODIVA and DIVA models. By combining these two frameworks, we can better understand how speech is planned, sequenced, initiated, executed, and controlled in real time.

To avoid overwhelming the reader with the complexity of the terminology, this section begins by outlining the core structure of the GODIVA and DIVA models in terms of

their overall functions (see Bohland et al., 2010; Civier et al., 2013; Meier & Guenther, 2023 for details). The technical components will then be illustrated with one example to clarify how they work together to support fluent speech.

Figure 3

GODIVA-DIVA Model of speech production



Note. Simplified schematic of the integrated GODIVA–DIVA model for speech production, adapted from Civier et al. (2013) and Guenther (2016).

The integrated GODIVA-DIVA model (see Figure 3) can be divided into three interconnected functional systems, each responsible for a key phase of the process of fluent speech production.

1. A PLANNING SYSTEM (GODIVA) that plans what to say and how to structure it. Two key representations are included:
   - A *structural frame representation* that encodes the prosodic structure of the utterance, including the number of syllables, their serial order, and stress pattern.
   - A *phonological content representation* that plans and selects the speech units to be produced.

Both representations include: (1) a *planning layer* that stores the speech units (phonological content) and their structural position (structural frame). A primacy gradient assigns higher activation to speech units that come first in the sequence, allowing the system to determine their production order. (2) A *choice layer* that selects, through a competitive selection, which speech unit is to be activated in which structural position.

2. A TIMING AND GATING SYSTEM that coordinates when the planned speech units are released for production and terminated. Two distinct loops are included:
    - A *Basal ganglia-Thalamus loop* (GODIVA) which ensures that the speech unit selected by the *phonological content choice layer* is released and sent to the motor execution system at the correct time. This release is triggered by the selection of a structural position in the structural frame choice layer.
    - A *Basal Ganglia-Premotor cortex loop* (GODIVA-DIVA) which oversees the initiation and termination of speech motor programs by monitoring the current sensory, cognitive and motor context, ensuring that selected motor commands are launched and completed at the appropriate moment, and thus move on to the next unit to be produced.
3. A MOTOR EXECUTION SYSTEM (DIVA) that transforms selected speech units into articulatory movements and monitors for production errors. This system includes:
    - A *Feedforward control system* that generates pre-learned motor commands associated with the selected speech units. Once the GODIVA system has selected a speech unit, it sends activation to the *plan layer* of the *speech sound map*, a syllabary that contains pre-learned motor programs for frequently used speech units, including the movement parameters and expected sensory outcomes (auditory and somatosensory). The motor program that receives the highest activation is selected in the *choice layer* of the *speech sound map*. The *initiation map*, monitored by the *basal ganglia–premotor cortex loop*, determines when a selected motor command should be released by monitoring the current sensory, cognitive and motor context, ensuring accurate speech timing. Once released, the motor program is sent to the *articulator map*, where it is translated into actual articulatory movements for speech production. After the speech movement has been adequately produced, the *articulator map* sends feedback to the *basal*

ganglia–premotor cortex loop, signaling the completion of the movement and allowing the system to terminate the current motor program and proceed to the production of the next speech unit.

- A *Feedback control system* that continuously monitors speech output for errors. It compares expected sensory feedback (from *auditory and somatosensory target maps*) with the actual sensory feedback received (*in auditory and somatosensory state maps*). Somatosensory feedback includes both proprioception, the internal sense of articulator position, movement trajectory, and coordination; and tactile sensation, which provides information about physical contact between articulators (such as lip closure for /p/ and the amount of force used). Auditory feedback monitors the acoustic characteristics of the produced speech sounds, ensuring they match internal auditory expectations. Discrepancies between expected and actual feedback are computed in *error maps*, and if mismatches are detected, corrective motor commands are issued via the *feedback control map* to the *articulator map*, allowing real-time adjustment during or after articulation.

Based on the GODIVA–DIVA model (see Figure 3), and assuming speech units are syllable-based, the production of "Lyra" proceeds as follows: Higher linguistic processes activate two *planning layers*: (1) a *structural frame* specifying two syllables with final-syllable stress, and (2) a *phonological content representation* buffering the syllables /li/ and /ʁa/, both organized by a primacy gradient favoring the first unit. Each planning layer communicates with its corresponding *choice layer*, selecting the first syllable and structural slot. The selected structural position triggers the *basal ganglia–thalamus loop*, which gates the release of /li/ into the motor system. The syllable /li/ activates its motor program in the *speech sound map*; if competing programs (e.g., /lo/) are briefly co-activated, /li/ wins selection. The *initiation map*, under control of the *basal ganglia–premotor cortex loop*, authorizes the release of /li/ at the appropriate moment. A feedforward command sends the motor plan to the *articulator map*, where it is executed. Completion of /li/ is signaled back to the *basal ganglia–premotor loop*, terminating the current syllable and authorizing the release of /ʁa/. Throughout, the *feedback control system* compares actual speech output to *auditory and somatosensory targets*. Detected mismatches trigger corrective commands to adjust articulation in real time.

Now that the two theoretical foundations have been presented, we will introduce seven major theories proposed in the literature to explain the physiopathology of stuttering, interpreted in light of these speech production models.

3. Major cognitive theories of stuttering

The literature on stuttering is rich with theories aiming to explain its nature and the mechanisms that underlie it, so rich, that one can easily feel overwhelmed by its density. This section aims to present the major cognitive theories that have been proposed, the empirical observations that prompted their development, and the key arguments put forth by their proponents. While a critical review of each of these theories lies beyond the scope of this thesis, the section concludes with a discussion of some key limitations that apply across all of them.

In this review, we have chosen to classify the major theories of stuttering into six domains: linguistic, auditory and audio-motor integration, somatosensory and somatosensory-motor integration, rhythmic and timing, motor, and multisensory integration. We acknowledge, however, that many theories overlap across these domains, and alternative classification schemes could also be justified.

These theories serve as the foundation for the cognitive battery developed in the present thesis to assess potential PROXIMAL CAUSES linked to stuttering (see Chapter 4). These theories are grounded in the models of fluent speech production detailed in the previous section, which provide the theoretical framework upon which their assumptions and mechanisms are built.

3.1 Evidence supporting the major theories of suttering

3.1.1 Stuttering as psycholinguistic disorder

As detailed in Levelt's (1989, 1999) model of speech production (see Section 2.1), even simple tasks such as naming an object require progressing through multiple, distinct levels of linguistic representation. A breakdown at any of these stages can disrupt speech production (Zhao & Lian, 2021).

Several lines of evidence have been put forward that suggest that stuttering may stem from dysfunction of one or several levels of linguistic representation (see Brocklehurst & Corley, 2011; Ntourou, Conture, & Lipsey, 2011; Packman et al., 2001; Seery et al., 2007, for review). For instance, studies have shown that PWS tend to stutter more frequently on grammatically complex utterances (Buhr & Zebrowski, 2009; Richels et al., 2010), on low-frequency words (Anderson, 2007), on function words rather than content words (Howell & Sackin, 1999; Richels et al., 2010), and at the beginning of sentences or words (Richels et al., 2010). Based on these findings, some researchers have speculated that stuttering may stem from a failure to efficiently and rapidly access one or more levels of speech planning. Empirical studies have further supported this view by identifying linguistic processing differences in PWS across multiple stages of planning. For example, some studies have reported slower semantic encoding (Bosshardt & Fransen, 1996; Maxfield, 2017), slower lexical encoding (Pellowski, 2011; Perkins, Kent, & Curlee, 1991), lexical selection difficulties (Newman & Ratner, 2007), and challenges resolving lexical competition (Maxfield, 2020, 2021). Other findings have pointed to slower syntactic encoding (Anderson & Conture, 2004; Tsiamtsiouris & Cairns, 2009), difficulties in phonological encoding (Postma & Kolk, 1993; Sasisekaran et al., 2006) and phonetic encoding difficulties (see Van Lieshout, 1995 for a review).

Based on this body of evidence, several linguistic theories have been proposed to explain how stuttering may emerge from disruptions in speech planning and language processing. These include the Fault Line Hypothesis (Wingate, 1988), the Neuropsycholinguistic Theory (Perkins, Kent, & Curlee, 1991), the Covert Repair Hypothesis (CHR; Postma & Kolk, 1993), the Vicious Circle Hypothesis (VCH; Vasic & Wijnen, 2005), the EXPLAN model (Howell & Au-Yeung, 2002), and the Variable Release Threshold Hypothesis (VRT; Brocklehurst, Lickley, & Corley, 2013). For example, Howell and Au-Yeung (2002) proposed that fluent speech is governed by two independent systems: planning (PLAN) and execution (EX), which must interact in a tightly coordinated manner. Stuttering is thought to arise when these two processes fail to synchronize properly, specifically, when the speech plan arrives too late for execution. This model implies that various dysfunctions in the planning process could lead to such asynchrony, resulting in disruptions to the flow of speech. On the other hand, the VCH (Vasic & Wijnen, 2005), an extension of the CHR hypothesis (see Postma & Kolk, 1993), posits that PWS exhibit

hypervigilant speech monitoring, leading them to detect minor or non-existent phonological encoding errors. This overactive monitoring results in excessive self-corrections and speech system breakdowns, contributing to stuttering-like behaviors. A critical prediction of this hypothesis is that stuttering-like disfluencies should decrease when PWS are engaged in distracting activities that divert attention away from speech monitoring, such as dual-task paradigm (see Eichorn, Pirutinsky, & Marton, 2019). Empirical studies support this claim: (1) Within these frameworks, the improvement of fluency observed with masked auditory feedback is interpreted as resulting from a reduction in self-monitoring demands, achieved by obscuring the auditory feedback of the speaker's own voice; (2) Tasks requiring PWS to speak while performing a secondary activity (e.g., finger tapping) reduce stuttering-like behaviors, as attention is shifted away from hypervigilant monitoring (Eichorn et al., 2016, 2019; Woumans et al., 2021).

3.1.2 Stuttering as an auditory/audio-motor integration disorder

As described in the GODIVA/DIVA framework (see Section 2.2), auditory processing plays a critical role in both the development, the planning and the ongoing control of speech production throughout the lifetime. In infancy, children begin constructing auditory models by first processing the speech of their caregivers, and later their own, thereby acquiring auditory representations of the speech sounds of their native language. These representations are gradually integrated with corresponding somatosensory and motor mappings of each speech sound. Once the speech motor system reaches maturity, auditory processing functions primarily as a feedback mechanism, monitoring discrepancies between expected and actual auditory input during speech production; but also plays a critical role in the timing of speech, as it is integrated within the brain's timing network (see Section 3.1.4). Two main types of acoustic cues contribute to auditory speech processing (Chen, 2011; Reinisch & Sjerps, 2011; Rosen, 1992; Shannon et al., 1995; Shetty, 2016): (1) spectral cues, which involve frequency-based information such as formant structure and fundamental frequency, are essential for identifying phonemes and distinguishing voice quality; (2) temporal cues, which include time-based features such as amplitude modulation, duration, and intensity, contribute to the perception of prosody, rhythm, and syllable boundaries.

Research has shown that temporal cues are especially crucial for parsing the rhythmic and prosodic structure of speech, which supports effective language comprehension (Chen,

2011). Disruptions in the processing of temporal cues can compromise auditory feedback mechanisms making it insufficiently clear (Fox et al., 2020; Kekade & Valame, 2014; Lotfi et al., 2020; Shetty, 2016). Such disruptions may lead to faulty auditory-integration representations, where mismatches arise between predicted and actual auditory input during speech production. These mismatches trigger error detection and compensatory corrections, which can interfere with fluent speech production (Chang et al., 2016; Lotfi et al., 2020; Max et al., 2004; Wieland et al., 2015).

Several lines of evidence suggest that PWS experience auditory processing difficulties. Differences have been observed in tasks that assess central auditory processing, including reduced performance in speech-in-noise recognition (Foundas et al., 2004; Shao, Bakhtiar, & Zhang, 2022), atypical patterns in dichotic listening tasks and altered right-ear advantage (Foundas et al., 2004; Robb, Lynn, & O'Beirne, 2013), and difficulties with temporal ordering tasks (Koca et al., 2024; Prestes et al., 2017). Additionally, the frequently reported improvements in fluency among PWS during conditions involving altered auditory feedback, such as masking, frequency shifting, or delayed auditory feedback, are, within this framework, consistent with a potential role of auditory processing factors in the manifestation of stuttering (see Lincoln, Packman, & Onslow, 2006 for review).

Based on the observation that PWS have poor auditory processing abilities, several researchers have proposed that it could lead to abnormal auditory-motor integration leading to stuttering-like behaviors (see Cai et al., 2012). Within the GODIVA/DIVA framework, four hypotheses have been advanced to explain how impairments in this integration could contribute to stuttering (see Bradshaw et al., 2021): (1) The sensory-motor modelling theory (Neilson & Neilson, 1987); (2) unstable or insufficiently activated internal models (Max, 2004); (3) overreliance on feedback control (Civier et al., 2010; Max, 2004); and (4) disrupted auditory prediction and feedback monitoring (Max & Daliri, 2019). According to the second hypothesis advanced by Max and collaborators (2004), stuttering may arise from the development of weak and unstable internal models, specifically, the feedforward commands that are responsible for predicting the sensory outcomes of speech motor commands. One important factor that can contribute to the formation of such faulty feedforward commands is poor auditory processing. When auditory processing is compromised, not only is the ability to establish accurate sensorimotor predictions reduced,

but the efficiency of processing auditory feedback during speech production is also diminished. Together, these difficulties can disrupt the flow of speech by increasing the likelihood of mismatches between the expected and actual sensory feedback.

3.1.3 Stuttering as a somatosensory/somatosensory-motor integration disorder

In line with auditory processing, somatosensory information is also crucial for both the planning and control of speech, as well as for maintaining accurate speech production throughout life (see Section 2.2 on the GODIVA/DIVA model). The somatosensory feedback system relies on two primary sources of information. The first source of somatosensory feedback is PROPRIOCEPTION which provides information about the position and movement of oral articulators, such as the lips, jaw, and tongue, independently of visual input. Sensory receptors located in the muscles and joints continuously transmit information about the status and movement of these structures to the brain. This feedback enables speakers to reliably monitor the configuration of their articulators during speech despite them not being visible. For instance, proprioception allows individuals to precisely control lip positions for different vowels: the lips are open for /a/, rounded for /u/, and stretched horizontally for /i/. Thus, by continuously informing the brain about the current configuration of the oral structures, proprioception helps guide the planning and execution of precise movements required for accurate and fluent speech production. The second source of somatosensory feedback is TACTILE PERCEPTION, or the sense of touch, which provides feedback from cutaneous receptors located in the oral mucosa, tongue, and lips. This sense allows speakers to detect contact between articulators (e.g., lips closing, tongue touching alveolar ridge) and to evaluate the amount of force applied during articulation. Tactile feedback is vital for producing sounds that involve articulatory contacts, such as when the lips are touching one another when producing the sound /p/ (see Guenther, 2016). To better illustrate how proprioception and tactile perception work together in speech, a dynamic example is provided below.

To accurately execute any speech sound, the system must first determine the current state of the articulators to generate motor commands (Parrell et al., 2019). This is done by integrating proprioceptive and tactile information. These sensory inputs allows the motor system to, then, precise the amplitude, direction, and timing of the required articulatory

movements. For instance, producing the speech sound /l/ in “Sélène” requires the tip of the tongue to move toward the alveolar ridge while the sides of the tongue are lowered to allow air to flow around them. To perform this movement accurately, the speech motor system must begin by assessing the current configuration of the articulators using proprioceptive and tactile feedback. In the case of “Sélène”, after the first syllable /s e/, the tongue tip is in a low or near-alveolar position—close to, but not actually touching, the alveolar ridge, as dictated by the articulatory demands of /s/ and /e/. To transition to /l/, the system relies on somatosensory input to determine this initial placement and then precisely plans the required articulatory adjustment: the tongue tip must rise and make firm contact with the alveolar ridge while the tongue sides are actively lowered to create the lateral channel characteristic of /l/. This sequence illustrates how the coordination and timing of lingual movement depend on continuously updated tactile and proprioceptive feedback. Once the movement is being executed, somatosensory feedback provides real-time updates on the position of the articulators (proprioception) and the contact force between them (tactile input). This feedback is then compared to the expected somatosensory pattern, allowing the system to detect and correct discrepancies even before the speech sound is fully articulated. It will also participate in the triggering of the next speech sound by informing where the articulators are, making somatosensory-motor integration an important part of the brain’s timing network (see Section 3.1.4).

Although the contribution of somatosensory deficits to stuttering has been less extensively studied than other domains, several lines of evidence suggest that PWS may experience somatosensory processing difficulties, particularly in relation to proprioception. As this topic constitutes the central focus of Chapter 2, we refer the reader to that section for a detailed review of empirical findings concerning proprioceptive impairments in stuttering. In contrast, relatively few empirical studies have addressed tactile perception in this population. Nevertheless, in line with the possibility of tactile processing difficulties, Fucci and collaborators (1985, 1991), as well as Petrosino and colleagues (1987), reported elevated vibrotactile detection thresholds on the tongue in individuals who stutter.

Although somatosensory processing has received less attention than auditory processing in stuttering research, no specific theories have yet been proposed that attribute stuttering primarily to somatosensory dysfunction. However, it is important to note that the

authors of the hypotheses described in the previous section acknowledge that disruptions in the motor control of speech, initially attributed to auditory processing difficulties, could also arise from impairments in somatosensory processing (see Chapter 4 and the General discussion).

3.1.4 Stuttering as a rhythmic/timing disorder

Timing is the essence of fluent speech. Speaking fluently requires the integration of multiple sources of information to ensure that speech sounds are initiated and terminated at the right moment (Moore, 2003; Schwartze & Kotz, 2024). Among these inputs are auditory (e.g., duration, stress pattern) and somatosensory (e.g., tongue position) context (Chang & Geuther, 2020; Moore, 2003). The *basal ganglia premotor cortex loop* (see Section 2.2 on the GODIVA/DIVA model) has been shown to integrate these inputs and to act as an internal clock, generating predictive rhythmic signals that help initiate and terminate speech sounds at the appropriate times (see Chang & Geuther, 2020; Moore, 2003; Schwartze & Kotz, 2024).

Several lines of evidence suggest that stuttering may stem from a timing disorder. For example, some PWS show difficulties with auditory rhythm perception (Wieland et al., 2015) as well as with producing rhythms in both verbal and non-verbal tasks (Cooper & Allen, 1977; Falk, Müller, & Dalla Bella, 2015; Sares et al., 2019; Slis et al., 2023). Neuroimaging studies also indicate atypical functioning in brain networks that support internal timing and rhythm in PWS (see Chang et al., 2016; Chang & Guenther, 2020; Etchell, Johnson & Sowman, 2014). Furthermore, according to the proponent of these theories, the observation that PWS often become more fluent when synchronizing their speech to external rhythmic cues, such as speaking with a metronome (see Frankford et al., 2022) or in unison with others ("choral" speech; see Park & Logan, 2015), provides additional support for a link between rhythm processing and stuttering (see Alm, 2004).

Based on this body of evidence, two main hypotheses have been proposed that implicate atypical rhythmic and timing processes in stuttering. The hypothesis of an inadequate neural clock for controlling speech segment duration (Cooper & Allen, 1977) suggests that disruptions in an "internal clock" responsible for controlling speech segment duration contribute to stuttering. Similarly, other accounts highlight the impaired ability of

brain networks supporting internal timing to generate appropriate timing cues (see Alm, 2004; Chang & Guenther, 2020 and Section 4.1.2).

3.1.5 Stuttering as a motor disorder

Once a speech unit has been successfully planned, selected, and initiated at the appropriate time, it must be transformed into articulatory movements for execution. This stage corresponds to the ARTICULATION level in Levelt's model, where the phonetic plan is realized as overt speech through motor execution. Similarly, within the GODIVA/DIVA framework, this process involves the activation of the ARTICULATOR MAP, which generates and coordinates the precise motor commands required for speech production (see Section 2). Like the execution of any motor skill such as walking, writing or driving, the smooth execution of an abstract speech motor plan requires the precise synchronization and coordination (in time and space) of muscle activity across the different subsystems involved in speech production; including the laryngeal, respiratory, and articulatory subsystems; at relatively high movement speed (Peters, Hulstijn, & Van Lieshout, 2000; Van Lieshout, 2017). Acquiring this high level of expertise necessitates progression from novice to expert, achieved through time, extensive practice, and focused attention, as well as a shift from explicit, conscious control to more implicit, automatic processes (Namasivayam & Van Lieshout, 2011). This progression can be observed during the acquisition of new skills, such as learning to drive a car, or in infants as they acquire complex motor abilities like walking or speech (Namasivayam & Van Lieshout, 2011).

Three lines of evidence suggest that stuttering may stem from difficulty in performing complex speech motor actions. First, studies have shown that PWS benefit less than PWNS from motor practice, both in achieving high levels of expertise in speech motor skills (Namasivayam & Van Lieshout, 2008) and in non-speech motor skills (Smits-Bandstra, De Nil, & Rochon, 2006), suggesting difficulty in shifting from a novice, implicit control of speech to a more expert, automatic control. Second, stuttering has been found to increase under conditions that destabilize the synchronization or coordination of movement by placing greater demands on the speech motor control system, for instance when increasing speech rate (see Smith & Kleinow, 2000; Smith et al., 2010). Third, conversely, stuttering tends to improve in situations that reduce the demands on the speech motor control system. According to this framework, many fluency-enhancing techniques, such as speaking in time with a

metronome or modifying auditory feedback, achieve this by decreasing speech rate, thereby facilitating the use of sensory feedback to guide motor control (see Namasivayam & Van Lieshout, 2011).

Based on this body of evidence, Namasivayam and Van Lieshout (2011) proposed the Speech Motor Skill (SMS) theory, which posits that stuttering arises from limitations in the speech motor control system's ability to perform complex speech motor actions. According to this view, PWS may be situated at the lower end of a speech motor skill continuum (Van Lieshout et al., 2014). The SMS theory further suggests that demands imposed by linguistic, motor, cognitive, or emotional factors can disrupt the precise synchronization and coordination of muscle activity required for fluent speech. In particular, movement amplitude, mediated through somatosensory feedback, is considered to play a critical role in the coordination and stabilization of speech motor control (Van Lieshout et al., 2014; Van Lieshout, 2017). From this perspective, larger movement amplitudes are thought to enhance somatosensory feedback and stabilize speech motor control, whereas smaller amplitudes may reduce coordination stability, increasing the risk of disfluency. According to proponents of this theory, this view is supported by empirical findings showing that reduced movement amplitude is associated with decreased coordination stability in speech gestures, independent of movement duration, and that such instability may underlie speech motor difficulties observed in stuttering (see Van Lieshout, 2017).

3.1.6 Stuttering as a multi-sensory-motor integration disorder

As discussed in Sections 2.2 and 3.1.2 to 3.1.4, sensory information plays a crucial role in the development, timing, and monitoring of speech. Although we have considered the importance of individual sensory modalities separately, it is important to recognize that these inputs are integrated into multisensory-motor representations in order to: (1) develop a sensorimotor representation of each speech unit; (2) determine the sensory context necessary to initiate and terminate the next speech segment in a sequence; and (3) monitor and correct speech errors in real time.

Several lines of evidence suggest that PWS might have multisensory-motor integration difficulty. Multisensory-motor integration can be evaluated using a finger-tapping task, in which participants are asked to synchronize their finger taps with an external

metronome beat. For precise synchronization, the brain must integrate somatosensory feedback from the finger (e.g., tactile contact, force application, and joint position) with auditory feedback from the metronome. Empirical studies have shown that PWS exhibit greater variability in inter-tap intervals (Sares et al., 2019; Slis et al., 2023; Van de Vorst & Gracco, 2017; Zelaznik et al., 1997). This increased variability suggests that sensory information is not effectively integrated, as their tapping rhythm fluctuates and often fails to align with the intended tempo. Furthermore, PWS tend to tap earlier relative to the beat than PWNS, demonstrating a larger Negative Mean Asynchrony (NMA) (Falk et al., 2015; Sares et al., 2019; Slis et al., 2023). According to the sensory accumulation theory (Aschersleben, 2002), NMA occurs because somatosensory feedback from finger tapping is processed more slowly than auditory feedback from the metronome. To compensate for this delay, the brain initiates the tap earlier to account for the slower somatosensory processing. A larger NMA in PWS implies that somatosensory feedback is processed more slowly relative to auditory feedback, leading to a mutlisensory-motor integration difficulties that could disrupt speech fluency. This observation is compatible with the often observation of fluency improvement resulting from delaying the auditory feedback of PWS (Kalinowski et al., 1996).

3.2 Challenges facing the major theories of stuttering

In the preceding section, we categorized major stuttering theories into six domains and outlined their empirical foundations. While these theories provide valuable insights, four critical limitations can be proposed that limit their explanatory power.

First, while empirical findings supporting various stuttering theories have been presented, it is critical to acknowledge that numerous studies report data inconsistent with theoretical predictions. This has contributed to a broader recognition that the stuttering literature is marked by inconsistencies and replication difficulties. Such variability across studies complicates efforts to build a cohesive understanding of the disorder.

For LINGUISTIC THEORIES, some studies fail to confirm the purported benefits of shifting attention away from speech (Lincoln et al., 2010; Oomen & Postma, 2002), and others even document increased stuttering-like behaviors under dual-task conditions (Bosshardt, 1999). Additionally, some research has not identified significant linguistic deficits or impairments in accessing planning system representations in PWS compared to

PWNS (see Hennessey, Nang, & Beilby, 2008). For AUDITORY/AUDIO-MOTOR INTEGRATION THEORIES, Assaneo et al. (2022) demonstrated that PWS, as a group, exhibit audio-motor integration comparable to PWNS, with only a subset of PWS showing atypical integration. Similarly, SOMATOSENSORY/SOMATOSENSORY-MOTOR INTEGRATION THEORIES face mixed evidence: while some studies report proprioceptive feedback integration deficits in PWS (Loucks & De Nil, 2006b), others fail to replicate these findings (Loucks & De Nil, 2012). RHYTHMIC/TIMING THEORIES are challenged by studies showing no timing deficits in PWS (e.g., Max & Yudman, 2003). For MOTOR THEORIES, Masapollo and collaborators (2021) failed to demonstrate that PWS learn new speech motor sequences differently than PWNS. Similarly, Korzeczek and collaborators (2020) found that PWS were able to retain finger sequences as well as PWNS. Likewise, MULTISENSORY-MOTOR INTEGRATION THEORIES are contradicted by research such as Max and Yudman (2003), which found no differences in variability during paced tapping tasks, and van de Vorst and Gracco (2017), who observed no increased NMA in PWS.

Second, the experimental data reported by proponents of the major stuttering theories predominantly reflect correlations or associations rather than direct evidence of causality. Nevertheless, a central goal of stuttering research, and thus of developing theoretical models, is to better understand the mechanisms underlying the disorder in order to inform effective interventions. As highlighted in recent reviews, most empirical findings in the stuttering literature are correlational, meaning they reveal relationships between variables (such as brain structure, timing performance, or linguistic abilities) and stuttering, but do not establish which factors are causal and which are consequential.

Determining causation is complex and context-dependent. In cognitive neuropsychology, Danks and Davis (2023, p.1), drawing on Woodward's (2005) interventionist framework, define causation as follows: "*C* causes *E* just when a possible (not necessarily feasible) intervention on *C* would probabilistically alter *E* when all else is held fixed". This approach emphasizes that to establish a causal relationship, one must consider whether manipulating the putative cause (*C*) would change the effect (*E*), assuming other factors remain constant. In practice, acquiring such causal knowledge is challenging, particularly in cognitive neuroscience and stuttering research. Empirical data often reflect correlations or associations between factors, which can be explained by multiple possible

causal pathways. According to Danks and Davis (2023) the gold standard for establishing causality is randomize manipulation of the suspected cause to observe what are the effects. However, in cognitive neuroscience, such experiments are rarely feasible: direct manipulation of neural factors (e.g., cortical networks) is often impossible, and randomization may not be practical when interventions depend on individual symptoms. Furthermore, when multiple potential causes are involved, as is common in complex and heterogeneous disorders like stuttering, the strategy of random manipulation does not scale well. These methodological constraints are compounded by the high degree of individual variability observed in stuttering, making causal inference even more difficult. Danks and Davis (2023) emphasize that robust background knowledge is critical for valid causal inference. When such knowledge is limited, datasets may contain numerous variables, but only subsets are causally relevant to the cognitive process or condition under investigation. Despite advances in stuttering research, persistent inconsistencies and replication challenges, likely stemming from the disorder's heterogeneity, constrain our foundational understanding of stuttering, thereby limiting the ability to draw definitive causal conclusions. Furthermore, even when causal explanations are proposed (e.g., mechanisms underlying the efficacy of fluency-enhancing techniques), alternative interpretations grounded in competing theoretical frameworks often remain equally plausible. For example, the beneficial effects of external timing cues (e.g., metronome-paced speech) can be attributed to rhythm/timing theories, multisensory-motor integration theories, or attentional frameworks, each offering a different account of the underlying mechanism. We propose that advancing our understanding of the cognitive heterogeneity underlying stuttering will enhance causal inference in this field. For example, large scale studies could identify subtypes where specific deficits (e.g., impaired somatosensory representation, impaired phonological encoding, timing difficulties) are predominant. Targeted interventions (e.g., sensory retraining for somatosensory-dominant subgroups) could then test whether modulating these factors directly improves fluency.

Third, a significant limitation arising from the predominant use of group-study paradigms in stuttering research is the restricted insight into how PWS perform across the various domain required for fluent speech (see Section 5). Most studies aim to test specific theories by identifying statistical differences in a limited number of tasks, often to confirm or refute a particular hypothesis. However, focusing on only a subset of abilities restricts our understanding of how PWS perform across the range of abilities required for fluent speech.

For example, testing the hypothesis that stuttering stems from an internal timing deficit, by assessing only rhythmic production tasks such as finger tapping, without evaluating sensory representations (e.g., auditory and somatosensory information) limits interpretation. Observed rhythmic difficulties could result from limited sensory representations that constrain sensory-motor integration, thereby impairing the timing network's ability to generate accurate timing cues. If sensory representation deficits are the primary issue in stuttering, then the problem extends beyond a purely rhythmic timing deficit, even though sensory processing is part of the broader timing network. This distinction implies that the core difficulty lies in sensory processing and integration, rather than in rhythm generation per se. Consequently, therapeutic interventions should focus on sensory training to enhance auditory and somatosensory feedback processing and integration, rather than solely on timing or rhythm training. Such sensory-focused therapies may better address the underlying sensory deficits, potentially leading to improved speech fluency. Similarly, linguistic theories that assess only one level of speech planning (e.g., lexical access) risk attributing the origin of stuttering to lexical retrieval deficits, when the underlying difficulty may actually originate from semantic processing issues that affect lexical access.

Fourth, the numerous theoretical perspectives proposed to explain stuttering, coupled with the observation of inconsistencies and difficulties in replicating empirical findings, suggest that stuttering may be a heterogeneous disorder (see Section 4 for further discussion). Even the slightest possibility that stuttering arises from multiple distinct mechanisms has major theoretical and methodological implications. From this perspective, we argue that group-study designs may be ill-suited for advancing our understanding of heterogeneous populations, as they risk obscuring important individual differences (see Section 5 for a more detailed discussion).

In this section, we have reviewed the major theoretical frameworks advanced in the stuttering literature. These theories form the foundation of the cognitive battery we designed to assess six domains implicated in fluent speech production (see Chapter 4). We have also proposed four critical limitations common to these frameworks. A key insight emerging from this critique is the likelihood that stuttering constitutes a heterogeneous disorder, reflecting variability in the underlying cognitive mechanisms across individuals. In the subsequent

section, we discuss in greater depth the arguments supporting the heterogeneity hypothesis of stuttering.

4. Stuttering or stutterings?

The notion that stuttering is a heterogeneous disorder is not new. In 1889, Ssikorski (p.44) already questioned this assumption, stating:

> The manifestations of stuttering are so diverse and complex that when dealing with this neurosis, one inevitably wonders whether these manifestations belong to a single specific disease or to multiple distinct ones, which, due to an inadequate analysis of the symptoms, have been grouped together under one category. The outward appearance of the disorder is so varied that it seems impossible to group together individual cases and compare them with each other.

However, it is important to clarify what is meant by *heterogeneity* in the context of this study. Individuals diagnosed with the same disorder may differ from one another on multiple levels. Drawing on the framework outlined in Section 1, these levels include: biological, cognitive, behavioral, and environmental. For instance, PWS may show differences in brain structure and function, as well as in genetic makeup. Environmental factors, such as stress, anxiety, or co-occurring conditions like attention deficits, may also influence the manifestation or severity of stuttering. Together, these variations may contribute to diverse behavioral expressions of stuttering across individuals, even if they share the same underlying cognitive mechanism. Conversely, it is also possible that some of the observed differences among PWS (such as their performance in tasks assessing cognitive behaviors, neurobiological organization, and responsiveness to treatments) may result from selective impairments in different components of the cognitive architecture that supports fluent speech. These differences at the cognitive level can exist even when individuals exhibit similar outward symptoms (see Thomas, 2003).

In the context of this study, and particularly in this section, we focus on the argument that stuttering may be underpinned by different underlying cognitive mechanisms (PROXIMAL CAUSES), independent of its biological origins or environmental influences (DISTAL CAUSES) (Sheikhbahaei, Millwater, & Maguire, 2023; St Onge, & Calvert, 1964). Even the slightest

possibility that stuttering arises from multiple distinct mechanisms has major implications for research (see Section 5) but has also major clinical implications (see Chapter 4 and General discussion).

4.1 Examining evidence of cognitive heterogeneity

Currently, there is no direct evidence for causal cognitive heterogeneity within the stuttering population. Consequently, it is not possible to assert that stuttering in one individual is caused by factor X and in another by factor Y. In this section, we outline four lines of evidence that suggest the presence of cognitive heterogeneity within the stuttering population. The first two constitute, in our view, direct evidence for such heterogeneity, while the remaining two represent more indirect or secondary indications.

4.1.1 Computational Modeling

Several works from Guenther's lab have directly simulated how stuttering-like behaviors can emerge by disrupting distinct neurocomputational processes.

Specifically, Civier et al. (2013) introduced two different types of impairment into their computational model (GODIVA), both affecting the thalamo-cortical circuit responsible for syllable selection and initiation. Although these impairments were biologically distinct, they each led to similar stuttering-like behaviors.

Earlier, in Civier et al. (2010), the authors showed that impairing the readout of feedforward motor commands forces the speech control system to rely excessively on feedback control (in their simulation, auditory feedback). This overreliance increases the detection of sensory errors, which, when they exceed a given error threshold, cause the system to reset and repeat the syllable, mimicking core stuttering behaviors (this support Max et al., 2004 theory).

Together, these studies reflect *direct proof* that stuttering can arise from diverse impairments, either at the level of movement initiation and sequencing (Civier et al., 2013) or from degraded feedforward command generation (Civier et al., 2010), highlighting the heterogeneity that can underlie the disorder.

4.1.2 Responses to treatments

Since the emergence of stuttering, considerable efforts have been made to develop effective treatments aimed at reducing stuttering-like behaviors. Two main lines of research have evolved: one focused on pharmacological interventions, and the other on cognitive and behavioral therapies.

The first reports of using pharmacological agents to treat stuttering date back to the 1950s (see Brady, 1991 for a review). Since then, a wide range of medications have been explored. Interestingly, the observation that different type of drugs, such as dopamine blockers and dopamine stimulants, can either alleviate or exacerbate stuttering symptoms has led researchers to hypothesize the existence of distinct neurochemical subtypes of stuttering exist (see Alm, 2004; Maguire et al., 2020; SheikhBahaei et al., 2023).

There are numerous non-pharmacological intervention options available for clinicians working with PWS. In a systematic review conducted by Baxter and collaborators (2015), five main categories of treatment approaches were identified: (1) feedback and technology-based interventions, which modify auditory feedback systems; (2) cognitive interventions, which aim to produce psychological change; (3) behavioral modification strategies, which target behavioral changes in the child, parent, or AWS; (4) speech motor interventions, which focus on modulating aspects of speech production such as respiration, phonation, or articulation; and (5) combined speech motor and cognitive interventions. The review, which included 112 studies, found that most reported positive effects for at least some participants, suggesting variability in treatment response. For instance, Ratynska and colleagues (2012) evaluated the effectiveness of a device that altered auditory feedback in 335 PWS. They observed a reduction in stuttering-like behaviors in 81% to 85% of participants, depending on the speaking context. However, 15% to 19% of participants showed no improvement or even exhibited increased difluency. These findings highlight the current challenge in the field: although many treatments options show promise, it remains unclear which individuals will benefit most from which intervention, pointing again to underlying variability in the mechanisms driving stuttering.

Once again, the observation that the same pharmacological or behavioral intervention can either exacerbate or ameliorate stuttering in different individuals provides *direct evidence*

that multiple, distinct mechanisms can underlie the disorder. Such divergent responses strongly support the presence of mechanistic heterogeneity within the stuttering population, highlighting the need to consider individual differences in underlying pathology when evaluating treatment effects.

4.1.3 Performance variability on cognitive processes

Group studies have reported that PWS show abnormalities across a wide range of abilities, including linguistic, somatosensory, auditory, rhythmic, both speech and non-speech motor functions, as well as executive functioning, memory, cognitive control, and even emotional and social cognition (see Section 3). But how should we interpret these findings? Do they imply that PWS experience deficits across all processes required for fluent speech? Could these differences reflect the influence of more general factors such as performance anxiety or co-occurring attention disorders? Or might they be the consequence of compensatory neural reorganization resulting from the persistent presence of stuttering during development? Or does it reflect the cognitive heterogeneity of stuttering? These possibilities, and their implications for understanding the cognitive mechanisms underlying stuttering, will be discussed in detail in Section 5.

There remains a critical need to directly demonstrate the existence of cognitive heterogeneity within the stuttering population, as well as to determine whether this heterogeneity reflects truly distinct causal mechanisms or merely represents the consequences and adaptations associated with the experience of living with stuttering. Distinguishing between inherent causal diversity and secondary effects is essential for a deeper understanding of stuttering.

4.1.4 Neurobiology

Several abnormalities in brain anatomy and function have been linked to stuttering. Importantly, the diversity of these neural differences may have distinct effects at the cognitive level, indicating that stuttering does not arise from a single underlying cause, but rather from multiple, distinct cognitive mechanisms. Below, we outline four arguments supporting this view.

First, in an effort to synthesize the neuroimaging literature on developmental stuttering, Etchell and collaborators (2018) conducted a systematic review of 111 studies. Their review revealed widespread structural and functional brain differences between PWS and their fluent peers. While the authors attribute the lack of consistency in specific regions and types of anomalies to methodological differences across studies, this inconsistency may also point to deeper explanations. Specifically, it could reflect (1) the existence of different subtypes of stuttering, each with distinct neural correlates; or (2) individual differences in neural reorganization that arise as compensatory responses to underlying cognitive difficulties. More recently, Matsuhashi and collaborators (2023) conducted a meta-analysis examining white matter integrity and functional connectivity across brain regions involved in fluent speech production. Their findings indicate that developmental stuttering is associated with both structural and functional alterations in multiple neural networks that support fluent speech and sequential motor processes.

Second, Chang and Guenther (2020) propose *three different loci* of impairment that involve the structure of the *Basal ganglia-premotor cortex loop* (the rhythm network) presented in the GODIVA/DIVA framework (see Section 2) that could contribute to stuttering. These biological impairments could have different repercussions in the cognitive level, suggesting different subtypes of stuttering. They discussed these theoretical perspectives in detail by reviewing empirical data.

- The first proposed locus involves dysfunction within the basal ganglia itself. This subcortical structure is crucial for any movement selection through two primary motor pathways: the direct pathway, which facilitates the activation of appropriate motor programs, and the indirect pathway, which suppresses competing motor programs. These processes are modulated by the neurotransmitter dopamine. To revisit the example given in Section 2.2.2: when the syllable /li/ is selected, its corresponding motor program is activated in the *speech sound map*, while co-activated alternatives like /lo/ must be suppressed. In fluent speakers, this suppression is effectively mediated by the indirect pathway. However, in PWS, elevated dopamine levels may impair the indirect pathway's function, leading to insufficient suppression of competing motor programs. As a result, the selection and stabilization of the correct motor plan are compromised. Chang and Guenther (2020)

suggest that this dysfunction could manifest in two ways: (1) a delay in the selection of the appropriate motor program, leading to speech blocks or prolongations, and (2) instability in the initiation signal, causing partial or failed initiation attempts that could result in repetitions. Support for this hypothesis comes from pharmacological findings: antipsychotic medications that block some dopamine receptors have been observed to reduce stuttering in some individuals, by improving the efficacy of the indirect pathway, thus the suppression of competing motor programs (Bothe et al., 2006).

- The second proposed locus involves impairments in projections from the cerebral cortex to the basal ganglia. These projections are essential because they convey the current sensorimotor and cognitive context required to generate rhythmic signals for the initiation and termination of speech motor programs. Civier et al. (2013) demonstrated through computer simulations using the GODIVA model that disruption of these projections can impair the timely release of planned speech units, potentially leading to stuttering-like behaviors. In support of this proposition, Chang & Guenther (2020) cite the work of Chang and collaborators (2016), which found that in typically fluent children, stronger activation of these projections was associated with better performance on a rhythm discrimination task, a task that relies on the precise processing of auditory temporal patterns. This correlation was absent in CWS, suggesting a deficit to process rhythm, a skill that is important for speech perception and production (Kotz & Schwartze, 2010).
- The third proposed locus of impairment lies within the cortical network responsible for implementing speech motor programs. For example, regions associated with the *speech sound map* in the DIVA model. This type of impairment corresponds to a dysfunction in the feedforward control system, which normally generates pre-learned motor commands for fluent speech. When this system is compromised, individuals may be forced to rely more heavily on the feedback control system. This system, though essential for error correction, is particularly sensitive to discrepancies between expected and actual sensory feedback (auditory or somatosensory). As a result, even subtle mismatches may trigger corrective processes, potentially disrupting fluent speech. Such overdependence on feedback may interfere with the generation of stable initiation signals by the basal ganglia,

ultimately contributing to the emergence of stutter-like behaviors. Additionally, Chang and Guenther (2020) suggest that the reduced activity often observed in the auditory cortex of AWS could reflect a compensatory mechanism, namely, a downregulation of auditory feedback monitoring to avoid triggering unnecessary corrections in response to minor production errors.

Third, acquired stuttering has been observed in association with a wide range of neurological conditions, including stroke, traumatic brain injury, seizure disorders, Parkinson's disease, and dementia (see Lundgren, Helm-Estabrooks, & Klein, 2010, for a review). Due to the diversity of conditions linked to acquired stuttering, it has become evident that numerous regions of the nervous system can give rise to stuttering symptoms. Lundgren and collaborators (2010) noted that acquired stuttering is not consistently associated with damage to any single discrete brain area. Rather, lesions implicated in acquired stuttering have been reported across various regions (Guenther, 2016), including all lobes of both cerebral hemispheres, the cerebellum, deep white matter, the thalamus, the brainstem, and the basal ganglia. Impairments in these different brain regions could affect distinct components of the cognitive system, potentially giving rise to various cognitive profiles of stuttering. However, more recently, Theys and collaborators (2024) analyzed lesion data and observed that, although the anatomical locations of stuttering-inducing lesions were heterogeneous, they all appear to be functionally connected to a common brain network. This network is centered around the left putamen, a structure within the basal ganglia known to play a central role in motor control. These findings suggest that disruptions to this core motor network, rather than damage to a specific region, may underlie the emergence of acquired stuttering.

Fourth, stuttering has also been described as an adverse reaction to a variety of pharmacological agents, including antiepileptics, antidepressants, antipsychotics, and methylphenidate (Ekhart et al., 2021; Nikvarz & Sabouri, 2022). These medications act on different neural systems, suggesting that stuttering-like behaviors can result from diverse alterations within the nervous system. Even more striking is the observation that some agents known to induce stuttering in certain individuals have been reported to reduce it in others. This paradox further supports the idea that multiple, potentially distinct, neural mechanisms can underline stuttering.

Here, we have presented four sources of evidence that suggest stuttering may arise from heterogeneous cognitive mechanisms. While we are not the first to consider this possibility, recognition of this heterogeneity has rarely progressed beyond proposals to classify stuttering into different subtypes, classifications that are typically not grounded in potentially distinct causal mechanisms. In the next section, we discuss the existing subtype proposals and examine the criteria on which they are based.

4.2 Classifying stuttering into subtypes

The heterogeneity of stuttering is widely acknowledged in the literature, and numerous efforts have been made to classify stuttering into distinct subtypes. These classification attempts have used a variety of criteria and can be broadly grouped into six levels (see Alm, 2005; Ambrose et al., 2015; Howell, 2011; Seery et al., 2007; Van Riper, 1982; Yairi, 2007):

1. Subtypes based on developmental course: these classifications distinguish stuttering according to onset (e.g., developmental vs. acquired) and outcome (e.g., recovery vs. persistence). While useful for longitudinal tracking, they offer limited insight into the cognitive underpinnings of the disorder and are therefore of limited clinical utility for diagnosis and treatment planning.

2. Subtypes based on core behaviors and severity: These classifications focus on observable differences in stuttering behaviors (e.g., repetitions, blocks, prolongations) and severity levels. However, they often lack a theoretical foundation and primarily offer descriptive rather than explanatory value.

3. Subtypes based on concomitant disorders: Some classifications are based on the presence of co-occurring conditions, such as language, motor, or attentional deficits. While these subtypes may suggest etiological diversity, the co-morbid features could also reflect overlapping or interacting conditions rather than variations in the core mechanisms of stuttering itself.

4. Subtypes based on etiology: These classifications propose different subtypes according to presumed causes. While some frameworks posit variants within a single etiological domain (e.g., psychological), others propose that entirely different

causes underlie each subtype. However, empirical support for many of these etiological distinctions remains limited or speculative.

5. Subtypes based on neurobiological factors: This approach focuses on biological differences such as genetic predispositions or structural and functional brain variations observed in PWS. Although these classifications acknowledge meaningful variation at the biological level, such findings are often difficult to interpret with respect to causality versus compensation (e.g., neural reorganization). Moreover, they do not clearly elucidate how these biological differences translate into distinct cognitive mechanisms underlying stuttering.

6. Subtypes based on treatment response: Some researchers have classified stuttering subtypes based on how individuals respond to pharmacological or non-pharmacological interventions. These classifications are particularly valuable, as they may reflect different causal mechanisms. When a specific treatment yields improvement for some but not all individuals, it suggests that the mechanism targeted by the intervention may be implicated in only a subset of cases, assuming that the lack of effect in other individuals cannot be attributed to alternative factors such as greater disorder severity or the presence of differing comorbidities. Thus, treatment response profiles may offer insight into the functional heterogeneity of stuttering and guide more individualized therapeutic approaches.

Despite growing recognition of the heterogeneity of stuttering, it has had little influence on the way stuttering studies are conducted, despite Alm's (2005, p.19) caution that "subtyping is likely to be a prerequisite for a detailed understanding of stuttering, as heterogeneous groups pose a significant challenge in research relying on group average data".

Despite intense research efforts that have significantly advanced our understanding of potential cognitive mechanisms underlying stuttering, further investigation is required to clarify the causal status of these deficits. A crucial step in this process is achieving a more comprehensive understanding of the cognitive heterogeneity present within the stuttering population. Accordingly, meaningful progress toward a full understanding of the disorder is likely to be slow unless future research designs explicitly account for the cognitive heterogeneity observed among people who stutter within the broad range of abilities that

sustain fluent speech. In the following section, we propose and outline a framework for systematically assessing this diversity in the study of stuttering.

5. Rethinking stuttering research

As described in Sections 3 and 4, there is inter-variability within the population of PWS, which has led several researchers to propose that the cognitive mechanisms underlying stuttering may be heterogeneous and that its etiologies are complex and multifactorial, with no single deficit sufficient to account for all cases. Despite this, recognition of heterogeneity has had little impact on research methodology, as noted by Yairi (2007, p.168):

> Among the large number of clinicians in the field and the smaller group of researchers, the inclination to treat or investigate stuttering as a single disorder, or "pathognomonic monolith" according to St. Onge and Calvert (1964, p. 160), has generally prevailed.

This implicit assumption of causal homogeneity, and the associated empirical search for a single, core deficit, has shaped the research methodology in the field of stuttering. Specifically, it has led to the widespread use of group-level analyses approches. In what follows, we discuss and critically examine the limitations of this approach. We then propose an alternative framework that, we argue, offers a more suitable path toward understanding the cognitive mechanisms underlying stuttering.

5.1 Limitations of group study designs in advancing stuttering research

Group-study designs aim to advance scientific understanding by identifying statistically significant differences in central tendency (such as the mean) between two or more groups, often using tasks or measures tailored to test specific hypotheses. Comparing the average performance can be highly effective for uncovering robust patterns and reducing the influence of random noise, especially when the individuals within each group are relatively similar with respect to the process under study. This approach has proven powerful in many domains. For example, in medical research, group-level analyses have demonstrated the public health benefits of widespread vaccination (Doherty et al., 2016). In such cases, the groups are well-defined by shared biological characteristics and a common response pattern.

A high degree of homogeneity is observed within the group, making group-level averages meaningful and informative.

However, what happens if there are reasons to believe that the group being investigated is composed of individuals who are not homogeneous? For example, what conclusions can be drawn from studying the physiology of zebras by mixing data from zebras and horses? Despite their surface similarities, these species differ in important aspects of their internal functioning. As a result, combining their data would inevitably lead to misleading conclusions regarding the physiological characteristics of either species, as true distinctions would be obscured and apparent "averages" might reflect properties that do not accurately describe either group. This illustrates the broader risk: when group-level analyses are applied to heterogeneous populations, critical differences can be masked, resulting in interpretations that neither capture the specific features of any subgroup nor reveal the mechanisms unique to each. The same parallel can be made at the cognitive level, if the members of a group differ in how they engage in cognitive processes, as it has been suggested for PWS, averaging their performance in cognitive assessments through group-study designs is likely to obscure meaningful individual differences and mask the existence of distinct cognitive profiles within the group. In such cases, the assumption of group homogeneity is violated, and group-level analyses may yield interpretations that do not accurately reflect the cognitive mechanisms present in any individual or subgroup, ultimately leading to misguided conclusions (Becker, 2023).

Thus, neglecting the possibility of cognitive heterogeneity in research designs has limited our understanding of stuttering and hindered progress toward identifying its true causes. To support this position, I will detail three specific points, regarding stuttering, in the following sections.

First, averaging the performance of individuals who differ in their abilities or impairments can lead to significant population sampling differences across studies (McCloskey & Caramazza, 1988; Schwartz & Dell, 2012; Towgood et al., 2009). This, in turn, contributes to the mixed and often inconsistent findings reported in the stuttering literature (Chang et al., 2025; Perez & Stoeckle, 2016; Yairi, 2007). For example (see Chapter 2), studies investigating a potential proprioceptive deficit in PWS, using a feedback perturbation paradigm, have yielded conflicting results. One study reported that, at the group

level, PWS were less efficient than controls in using proprioceptive feedback to make precise jaw movement (Loucks & De Nil, 2006a), whereas a subsequent study found no significant group differences under similar conditions (Loucks & De Nil, 2012).

Second, statistical artifacts that results from group-level analyses can mask or overestimate meaningful possible causal explanation. For instance, group differences may appear even when only a subset (or none) of the individuals in the stuttering group actually present the deficit, leading to an OVERESTIMATION of its prevalence (see De Nil & Abbs, 1991; Van Thorre et al., 2025). Conversely, both groups may appear to perform similarly, while individual data reveal that some PWS do exhibit marked difficulties, thereby MASKING a potential underlying impairment (see Daliri et al., 2013). The key risk here is falling into what is described as an "averaging artifact" (Shallice & Evans, 1978) where group-level measures may fail to reflect the performance of any individual in the group. Such aggregation can obscure meaningful individual differences and misrepresent the underlying cognitive mechanisms (Becker, 2023; Fedorenko, 2021; McCloskey & Caramazza, 1988; Towgood et al., 2009; Shallice & Evans, 1978; Schwartz & Dell, 2012; Williams, 2010).

Third, group-based studies have shown that PWS exhibit difficulties across nearly all cognitive domains involved in fluent speech. But how should we interpret this broad pattern? Several possibilities must be considered: (1) These difficulties may arise from general, non-specific factors such as performance anxiety, attentional deficits, or broader neurological influences (Anderson & Ofoe, 2019; Angelidis et al., 2019). Notably, PWS often report high levels of anxiety (Bernard et al., 2022; Iverach et al., 2016; Smith et al., 2014). (2) The observed impairments may reflect secondary effects of stuttering on cognitive development (Packman et al., 2007; Thomas & Karmiloff-Smith, 2002). (3) The cognitive basis of stuttering might still be causally homogeneous, with variability arising from unrelated individual differences or measurement noise, for instance, co-occurring disorders that vary across PWS (Thomas, 2003). (4) Alternatively, stuttering may reflect a truly heterogeneous condition, with multiple underlying causes, each producing its own consistent behavioral profile. When aggregated, these profiles are superimposed and obscured at the group level (Thomas, 2003).

The critical question, then, is: *How can we distinguish between these possibilities?* Group study designs, which typically rely on a small number of tasks, provide limited insight

into how each PWS performs across the range of cognitive processes involved in fluent speech. Access to detailed individual profiles across multiple tasks could help disentangle the nature of the variability and guide us toward a more precise understanding of the disorder. For example, such data could reveal three possible scenarios: (1) profiles showing deficits across all domains, which would be difficult to interpret and might suggest that stuttering is related to general, non-specific cognitive factors, or reflect the effect of stuttering on cognitive development; (2) a single (or a combination of) domain consistently impaired in all PWS (NECESSARY CONDITION), and the only deficit in some PWS (SUFFICIENT CONDITION), will suggest a homogeneous origin with individual variability; or (3) distinct (clear) profiles with different impairments across individuals, indicating the presence of multiple, heterogeneous cognitive etiologies.

To address these possibilities, we propose that a neuropsychological single-case design provides an effective methodological framework.

5.2 Advancing stuttering research through single-case methodology

Originally developed within the field of cognitive neuropsychology, the single-case approach has played a crucial role in improving our understanding of the structure of cognitive systems, and above all in shedding light on the cognitive underpinnings of various disorders. Before examining how single-case designs can contribute to stuttering research, it is helpful to first outline the core assumptions upon which this approach is based. These assumptions provide the theoretical foundation for using detailed individual investigations to identify and interpret cognitive impairments. We will then illustrate its effectiveness in the study of acquired disorders, focusing specifically on language impairments such as word-finding difficulties resulting from brain injury. Finally, before considering its application to stuttering, we will address how it has informed our understanding of several such conditions.

5.2.1 How the single-case approach can improve our understanding of the cognitive basis of disorders

When studying human cognition, whether in typical or pathological states, certain assumptions must be made in order to interpret the observation. In this thesis, we adopt the

four assumptions proposed by Caramazza (1984, 1986), which provide a theoretical framework for analyzing and interpreting participants' performance profiles. Originally formulated to allow generalization from single-case studies of brain-damaged individuals to the cognitive architecture of healthy individuals, we argue that they are equally relevant for understanding atypical cognition more broadly. This framework enables us not only to make inferences about the structure of the typical cognitive system, but also to identify the specific mechanisms that are impaired in atypical cases. For example, it has helped reveal that similar symptoms, such as word finding difficulties, can result from disruptions to different cognitive components, thus emphasizing the value of detailed case-level analysis for uncovering heterogeneous pathways to similar clinical manifestations (see Section 5.2.2).

The first one is the UNIVERSALITY ASSUMPTION, which posits that individuals share a common cognitive architecture, making inferences from single cases generalizable to the broader population. In other words, although individuals may differ in their developmental trajectories or neurological outcomes, they are assumed to rely on the same underlying cognitive system. Caramazza and Coltheart (2006) emphasize that this assumption applies across all areas of cognitive psychology, enabling researchers to generalize findings, whether derived from single-case studies or group investigations, across both typical and atypical populations. Failing to recognize the principle of universality would severely limit our ability to construct and evaluate models of cognitive processing, thereby undermining our understanding of both typical and pathological functioning. As emphasized by Tempel (1997), rejecting the universality assumption would impose significant constraints on our ability to advance our understanding of human cognition:

> If one were to admit individual differences in the wiring diagrams of the functional architectures themselves, such that some brains had a fundamental organizational structure that differed from others, then the enterprise of cognitive neuropsychology applied to adults or children could be under threat (Temple, 1990a). If there were potentially different architectures, then constraints on the models that may be developed disappear. There are too many degrees of freedom. The models become untestable and irrefutable because any pattern of data could be fitted with the postulation of a further variant in architecture.

It is therefore essential to ground experimental designs in well-specified models of the cognitive architecture being investigated, while remaining open to revising these models in light of empirical discoveries. On a more philosophical note, if we cannot assume that cognitive processes are organized similarly across individuals, then what becomes of the distinction between typical and atypical cognition, particularly when referring to different developmental trajectories or neurological conditions? Without a shared reference framework, the very notion of 'atypical' loses its meaning, as it implies deviation from a normative cognitive architecture that do not exist.

The second one is the FRACTIONATION ASSUMPTION, which posits that a lesion can selectively impair a single component of a cognitive process while leaving others intact. This assumption implies that complex cognitive functions are made up of separable components or modules. This leads to the two following, more basic assumptions: the MODULARITY AND TRANSPARENCY ASSUMPTIONS.

The MODULARITY ASSUMPTION is the idea that the cognitive system consists of functionally distinct, basic operating modules (Fodor, 1983). For example, naming an object involves access to different levels of representation that can be selectively impaired. One of the strongest arguments in favor of this assumption is the existence of dissociations (contrasting normal performance on one task and impaired performance on another one) and more particularly double dissociations in single-case studies. Double dissociations occur when two patients exhibit complementary patterns of impairment: for example, Patient A shows impaired performance on Component 1 but preserved performance on Component 2, whereas Patient B demonstrates the opposite pattern, intact performance on Component 1 but impaired performance on Component 2. Such findings provide strong evidence that the two components operate independently, supporting the idea of modular organization in the cognitive system.

The TRANSPARENCY (OR SUBTRACTIVITY) ASSUMPTION posits that, after neurological damage or atypical development, an individual's cognitive system reflects the intact architecture of a healthy system minus any impaired modules. While the brain may reorganize to compensate for damage, such reorganization is limited to redistributing existing functions among intact modules; it cannot add entirely new modules. This ensures that the

cognitive architecture following damage remains fundamentally comparable to the healthy system, allowing valid inferences about normal cognition from studies of impairment.

Although these assumptions are not directly testable, proponents of the approach emphasize that they should not be taken for granted and may ultimately prove to be right or wrong. Coltheart (2017) argues that, despite this, they remain defensible even 30 years after their original formulation, citing Ellis and Young (1988, p.16) to support the view that such assumptions are necessary for meaningful progress in cognitive neuropsychology:

> The philosopher of science Imre Lakatos has argued that every science has at its core a set of assumptions which are not directly testable… These assumptions may be right, or they may be wrong – the only way that the scientist working in a particular area will know whether its assumptions are right or wrong is by seeing whether the whole approach advances or flounders.

This approach has proven effective in advancing our understanding of nearly all domains of cognition (Towgood et al., 2009), for example in explaining how the brain processes language based on brain damaged patients (see section 5.2.2). This logic has been extended to the study of developmental disorders (see section 5.2.3). Such designs help determine whether shared behavioral diagnoses (e.g., dyslexia) reflect a single underlying cause or multiple heterogeneous cognitive etiologies (but see Bishop, 1997; Karmiloff-Smith, 1997, 1998; Thomas & Karmiloff-Smith, 2002, 2009 for a discussion about the use of the cognitive neuropsychological approach for developmental disorders). Proponents of the approach (see Caramazza, 1986) notably criticized group-based designs, pointing out that individuals with similar symptoms may have different underlying cognitive deficits. Aggregating their data can result in average profiles that reflect none of the individual cases, obscuring true cognitive mechanisms. In contrast, single-case designs allow researchers to identify dissociations and target specific components by mapping the individual's strengths and weaknesses. This supports the development, testing, and refinement of cognitive models of typical functioning (Rapp & Caramazza, 2002), but it is also valuable for improving our understanding of the causes of symptoms, and for clinically oriented work, as clinical work is inherently focused on determining the specific impairments and needs of each patient, rather than identifying group-level trends (Towgood et al., 2009).

In the next section, we will show how the single-case approach has proven valuable in advancing research of brain damaged patient by illustrating its impact on the study of acquired disorders, focusing on language impairments (more specifically, word finding difficulties) resulting from brain damage.

5.2.2 Benefits of the single-case approach in acquired disorders

Even a seemingly simple task like naming an object engages several distinct levels of cognitive processing, each of which can be selectively disrupted by brain lesions. Anomia, or word-finding difficulty, is a common symptom after such damage; however, cognitive neuropsychological investigations—such as those of KE, RGB, and HW (Caramazza & Hillis, 1990; see also Caramazza, 2000; Hillis et al., 1990; Kay & Ellis, 1987), have demonstrated that anomia can arise from impairments at different stages of word retrieval. For instance, KE's deficit was traced to impaired access to semantic information, while RGB and HW exhibited post-semantic or phonological encoding impairments. Careful analysis of patients' errors across tailored tasks allows researchers to work backward from symptoms to map which components are impaired and which are preserved, based on theoretical models of language processing.

Cognitive neuropsychological research on anomia has achieved two major advances: it has significantly increased our understanding of how word retrieval is organized in the brain—particularly the processes underlying lexical access—and, at the same time, it has demonstrated that a common symptom like anomia can arise from different underlying cognitive mechanisms. This insight has guided clinical practice, enabling clinicians to pinpoint the specific locus of a patient's anomia and to select more effective, individualized treatments. To give a more practical example, a blocked nose is a common symptom of various conditions (e.g., cold, flu, allergies, sinusitis, nasal septum deviation, nasal polyps). In order to treat their patients effectively, physicians must conduct a thorough medical examination to determine the underlying cause and propose the most appropriate treatment. For instance, no responsible physician would propose surgery to correct a deviated septum in a patient whose nasal congestion is actually caused by allergies.

So far, we have demonstrated the benefits of the single-case approach in investigating acquired disorders, especially in brain-damaged patients. This method has also been

successfully applied to the study of developmental disorders, yielding valuable insights into the cognitive mechanisms that underlie conditions such as, dyslexia, and dyscalculia.

5.2.3 Benefits of the single-case approach in developmental disorders

Developmental dyscalculia and dyslexia are two specific learning disorders, defined (behaviorally) respectively as difficulties in acquiring mathematical and reading competencies despite adequate instruction, that share a similar research history with stuttering in terms of methodological approaches. Like stuttering, both dyscalculia and dyslexia have been extensively studied, primarily through group-based designs. Yet, despite a substantial body of research, their underlying cognitive mechanisms remain debated. Similarly to stuttering, both disorders are marked by significant cognitive heterogeneity, suggesting that different individuals may exhibit different underlying cognitive deficits. To better understand the heterogeneity observed in these disorders, and potentially differentiate between the competing theoretical hypotheses put forward by group studies, several research teams have conducted multiple single-case investigations (see De Visscher & Noël, 2013; Ramus et al., 2003; Träff et al., 2016).

Concerning dyslexia, Ramus and collaborators (2003) administered a battery of tasks designed to test three major hypotheses proposed in the literature: the magnocellular hypothesis, the phonological theory, and the cerebellar theory. Sixteen individuals with dyslexia and sixteen typically developing controls participated in the study. By analyzing individual performance profiles, the authors identified a range of patterns among the dyslexic participants. Notably, they found that a phonological deficit was present in all individuals with dyslexia, suggesting that it may represent a NECESSARY CONDITION for the disorder. Moreover, for some participants, it was the only observed deficit, no auditory, motor, or visual impairments were detected, indicating that it may also serve as a SUFFICIENT CONDITION. These findings have both theoretical and clinical implications: they support the phonological theory of dyslexia while also acknowledging the presence of additional sensory and motor difficulties that vary across individuals and may exacerbate the severity of reading difficulties. This underscores the importance of conducting in-depth assessments of both research participants and clinical patients to determine how such additional factors may contribute to individual profiles of impairment. Furthermore, the causal role of the

phonological impairment must be assessed on a case-by-case basis, ideally through intervention studies designed to determine whether targeted phonological training leads to measurable improvements in reading ability.

Concerning dyscalculia, Träff and collaborators (2016) administered a comprehensive battery of tasks designed to assess a wide range of cognitive skills required for mathematical learning, based on the major theoretical accounts proposed in the literature. Four children diagnosed with developmental dyscalculia participated in the study and were compared to different normative groups. By analyzing individual performance profiles, the authors found that although all four children exhibited similar mathematical difficulties, they demonstrated considerable variability regarding their cognitive difficulties and thus their cognitive profiles. These distinct cognitive profiles suggest the existence of multiple subtypes of developmental dyscalculia and support the view that a group-study approach may not be the most appropriate, given that individuals with dyscalculia do not appear to form a homogeneous group in terms of the cognitive abilities underlying mathematical learning.

After presenting two examples illustrating how the cognitive neuropsychological approach can enhance our understanding of variability in developmental disorders, the next section will outline how we are going to implement it in the study of stuttering.

5.2.4 Implementation of the single-case approach to stuttering

Given the possibility that stuttering may stem from cognitive heterogeneous origins (see Section 4), a multiple single-case study design appears particularly well-suited for its investigation (see Chapter 3 and 4 for implementation details). In this thesis, we adopt the causal modeling framework proposed by Jackson and Coltheart (2013), adapted for the study of stuttering (see Section 1), with a specific focus on identifying the PROXIMAL CAUSES of stuttering, independently of DISTAL CAUSES.

This approach aims to shed light on cognitive heterogeneity by examining detailed profiles of strengths and weaknesses in PWS, across cognitive domains implicated in fluent speech production. These domains are defined based on the models of speech production presented in Section 2, and five major theoretical accounts of stuttering: psycholinguistic, auditory, somatosensory, rhythmic, and motor (see Section 3). To this end, we developed a comprehensive cognitive battery designed to assess each of these domains.

By analyzing the performance of several PWS, we aim to determine whether stuttering results from homogeneous or heterogeneous mechanisms. Specifically, we examine whether any observed deficits can be considered NECESSARY CONDITIONS (present in all individuals) and/or SUFFICIENT CONDITIONS (the only detectable deficit in some individuals) for stuttering.

6. Research question

Through a comprehensive assessment of the cognitive abilities underlying fluent speech, this thesis aims to clarify the patterns of strengths and weaknesses experienced by PWS within these domains. Based on the literature review and theoretical considerations presented above, it is hypothesized that these cognitive patterns will vary across individuals, thereby offering the first empirical demonstration of cognitive heterogeneity among AWS.

7. Outline of the thesis

This thesis is composed of three chapters and concludes with a chapter dedicated to a general discussion of the results. Below is a brief overview of the content and structure of each chapter.

Chapter 2 stems from the observation that while the potential contribution of auditory processing to stuttering has received considerable attention in the literature, somatosensory processing—despite being equally critical for the control of fluent speech—has not been investigated nearly as extensively. Accordingly, Chapter 2 focuses on reviewing and critically evaluating the existing scientific evidence relevant to the hypothesis that PWS may exhibit a proprioceptive deficit. This section provides an integrated synthesis of behavioral and neurofunctional findings related to proprioceptive processing in PWS, drawing on data from studies of oral stereognosis, assessments of proprioceptive acuity, and investigations involving perturbations of proprioceptive feedback during speech.

Chapter 3 aims to determine whether PWS exhibit specific weaknesses in phonological working memory, that is, within processes independent of speech production. By critically examining previous findings, we outlined four methodological limitations of prior research and proposed a novel experimental design that enables a clearer investigation of this question, incorporating both group-level and single-case approaches.

Chapter 4 will implement the multiple single-case approach, borrowed from cognitive neuropsychology, to explore the cognitive underpinnings of stuttering. Through this approach, each participant will undergo a series of experimental tasks that assess diverse cognitive domains implicated in fluent speech, such as linguistic, auditory, somatosensory, rhythmic, and motor abilities. Analysis of the cognitive profiles of each participant will enable the identification and mapping of distinctive patterns of strengths and weaknesses within the domain investigated. By focusing on individual data, Chapter 4 seeks to empirically demonstrate the existence of cognitive heterogeneity inside the stuttering population and to reflect on potential contributing factors that may be linked to the emergence of suttering.

Finally, this thesis will conclude with a General Discussion section. The main findings will be summarized and interpreted in relation to existing literature. The implications of these results will be considered for both clinical practice and future research. We will also acknowledge the limitations of the study and outline perspectives for further investigation. The thesis will close with a general conclusion.

Chapter 2. Bégaiement et proprioception

Abstract

Les personnes qui bégaient seraient anormalement dépendantes de l'information auditive pour contrôler la production de leur parole. Cette dépendance excessive pourrait s'expliquer par un trouble proprioceptif. Ce chapitre fait le point sur les données scientifiques existantes pertinentes par rapport à cette question : les personnes qui bégaient ont-elles un déficit proprioceptif ?

Référence

Van Thorre, A., Vannuscorps, G. (2018). Bégaiement et proprioception. In B. Pierart (Ed.), *Neuropsychologie du bégaiement : de la compréhension au traitement* (pp. 127- 136). Mardaga

1. Introduction

De nombreuses observations suggèrent que les personnes qui bégaient sont anormalement dépendantes de l'information auditive pour contrôler la production de leur parole (Max et al., 2004 ; Tourville, Reilly, & Guenther, 2008). Le contrôle auditif de la parole n'est pourtant pas le plus efficace. En effet, l'information auditive en provenance de notre voix est lente : il faut environs 100 à 150 millisecondes pour qu'elle atteigne notre système auditif (Guenther, 2016 ; Marslen-Wilson & Tyler, 1981). Cette latence ne permet au système de contrôle que de corriger d'éventuelles erreurs après qu'elles se soient produites provoquant un « redémarrage » de la parole, des répétitions. Pour cette raison, cette dépendance au contrôle auditif est considérée par certains auteurs une origine causale possible du bégaiement (Civier et al., 2010). En faveur de cette hypothèse, il a été observé que les personnes qui bégaient bénéficient fortement de différentes techniques qui permettent de pallier cette dépendance, par exemple le masquage auditif (Sutton & Chase, 1961 ; Wingate, 1970), la modification du feedback auditif (Kalinowski et a l., 1993; Macleod et al., 1995) ou la réduction de la vitesse de parole (Andrew et al., 1982). A titre d'anecdote, dans son livre « The Nature of Stuttering » Van Riper (1982) décrit une personne dont le bégaiement a disparu soudainement « après un accident qui a rendu cette personne complètement sourde. »

Cependant, s'il est aujourd'hui admis par la plupart des chercheurs que les personnes qui bégaient dépendent davantage que la normale des informations auditives, l'origine de cette dépendance reste peu claire. Selon certains auteurs, il s'agirait de la conséquence d'une faiblesse se situant au niveau des commandes motrices apprises durant le développement normal ou de leur interprétation par le système moteur (Alm, 2004 ; Civier et al., 2013). Pour d'autres auteurs, il s'agirait de la conséquence d'un système de contrôle hyperactif de la parole (Vasic, & Wijnen, 2005). Cependant, une autre possibilité est que la dépendance au feedback auditif est la conséquence d'un trouble proprioceptif. La proprioception est le sens qui nous permet de connaître la position des différentes parties de notre corps sans devoir les observer. C'est le sens qui vous permet de savoir, par exemple, si vos lèvres sont actuellement ouvertes ou fermées, si votre langue est relâchée ou contractée. Dans le développement normal, le contrôle de la parole passe progressivement d'un contrôle auditif qui est lent et imprécis, à un contrôle proprioceptif lorsque les associations entre les commandes motrices

et leurs corrélats proprioceptifs sont stabilisées. Ce passage au contrôle proprioceptif est crucial car, étant beaucoup plus rapide, il permet de s'assurer que le son qui va être produit correspond aux intentions du locuteur *avant* que le son ne soit produit (sur base de la position des articulateurs). Un trouble proprioceptif pourrait donc empêcher ou forcer l'utilisation des feedbacks auditifs. Cette hypothèse n'est pas nouvelle. Van Riper (1971), écrivait déjà : « le contrôle auditif, important lors des premières productions langagières de l'enfant, laisserait, lors du développement, la place au contrôle kinesthésique et proprioceptif ; ce transfert serait imparfait chez le bègue. ». Dans ce chapitre, nous faisons le point sur les données existantes pertinentes par rapport à cette question : est-ce que les personnes qui bégaient ont un déficit proprioceptif ?

2. La proprioception dans le bégaiement

2.1 Les données comportementales

Jusqu'à présent, onze études ont étudié les capacités proprioceptives orofaciales de personnes qui bégaient (PQB). Ces études peuvent être réparties en trois catégories en fonction de la méthode d'évaluation utilisée : la reconnaissance de forme, l'acuité proprioceptive, et la perturbation du feedback proprioceptif.

2.1.1 Les études portant sur la reconnaissance de formes

Dans trois études, les capacités proprioceptives des PQB ont été testées par le biais d'une tâche dans laquelle il leur était demandé de reconnaitre des formes géométriques fixées sur un support, attachées à une ficelle, et placées dans leur cavité buccale (Jensen et al., 1975 ; Martin et al., 1981 ; Stewart et al., 1985). Dans l'étude de Jensen et collaborateurs (1975), les sujets (30 adultes qui bégaient et 30 adultes contrôles) devaient identifier, parmi 10 formes géométriques, quelle était celle qu'ils avaient en bouche en la désignant sur une feuille illustrant l'ensemble des formes ainsi que deux distracteurs. Dans les études de Martin et al. (1981) sur 20 adultes qui bégaient et 20 adultes contrôles, et de Stewart et al. (1985) sur 16 enfants qui bégaient et 16 enfants contrôles, les sujets devaient décider si deux formes géométriques placées dans leur bouche à quelques secondes d'intervalle étaient identiques ou différentes. Seule l'étude de Martin et al. (1981) a rapporté des différences significatives entre les performances moyennes des deux groupes. Notons néanmoins que dans les deux

autres études, les PQB obtenaient en moyenne un score moins élevé que les sujets contrôles (score moyen entre 2.7% et 8% moins élevé) et certaines des PQB faisaient plus d'erreurs que le moins bon des sujets contrôles. Par exemple, dans l'étude de Stewart et al. (1985) – la seule étude pour laquelle nous disposons des données individuelles de chaque participant – 25% des enfants qui bégaient (4/16) ont commis plus d'erreurs (de 16 à 50%) que le moins bon des sujets contrôles. Dans ce contexte, l'absence de différence statistiquement significative entre les moyennes de groupes semble être la conséquence d'une hétérogénéité importante des capacités proprioceptives au sein du groupe de personnes qui bégaient. Cette interprétation est étayée par l'observation d'une dispersion plus large des performances des PQB que des sujets contrôles dans les trois études.

2.1.2 Les études portant sur l'acuité proprioceptive

Dans sept études, les capacités proprioceptives des PQB ont été testées par le biais d'une tâche dans laquelle il leur était demander de réaliser un mouvement de la sphère orofaciale (le plus souvent de la mandibule) le plus précis possible (par exemple, faire le plus petit mouvement possible) avec ou sans la possibilité d'observer leurs mouvements (feedback visuel). Le raisonnement des chercheurs était le suivant : si les personnes qui bégaient ont un trouble proprioceptif, alors, elles devraient exécuter des mouvements moins précis que ceux des sujets de contrôles lorsqu'ils ne bénéficient que d'information proprioceptive (ils ne peuvent pas observer leur mouvements) mais aucune différence ne devrait être observée (ou une différence plus petite) lorsqu'ils bénéficient de l'information visuelle.

Dans une *première étude*, les participants (6 adultes qui bégaient et 6 adultes contrôles) devaient réaliser le plus petit mouvement possible vers le haut ou vers le bas avec leur mandibule, leur lèvre inférieure, leur langue ou leur index avec ou sans feedback visuel (De Nil, & Abb, 1991). Aucune différence de précision n'a été observée entre les deux groupes lorsque les participants bénéficiaient du feedback visuel. En l'absence de feedback visuel, par contre, les mouvements des articulateurs (mais pas de l'index) des PQB étaient significativement moins précis (i.e., plus large). Une *deuxième étude* a répliqué ces résultats (Archibald, & De Nil, 1999). Donc, il semble que les PQB soient moins précis pour réaliser un mouvement le plus petit possible de leurs articulateurs lorsque la seule information disponible pour le contrôle moteur est l'information proprioceptive.

Aucune différence statistiquement significative entre les performances moyennes des deux groupes n'a été mise en évidence dans une *troisième étude* dans laquelle les sujets devaient ouvrir leur bouche d'une amplitude précise (18 mm) sans feedback visuel (Loucks, & De Nil, 2006a). Cependant, cette absence de différence est probablement la conséquence d'un manque de sensibilité de cette tâche. En effet, d*eux autres études* qui ont utilisé un paradigme similaire ont observé que les PQB réalisaient des mouvements significativement moins précis (plus amples et plus variables) que les sujets contrôles lorsqu'ils devaient les effectuer en se basant uniquement sur les informations proprioceptives (condition non visuelle) et le plus rapidement possible (Loucks, & De Nil, 2006b, 2012). Dans une *sixième étude* les participants devaient prononcer la voyelle /a/ en ouvrant leur bouche d'une amplitude précise (soit 6 mm, soit 15mm) soit lentement, soit deux fois plus rapidement, en ne disposant d'aucun feedback (visuel ou auditif) autre que proprioceptif (Loucks, De Nil, & Sasisekaran, 2007). Encore une fois, dans cette tâche, les PQB étaient significativement moins précises (mouvements plus amples et plus variables) que les sujets contrôles. Dans une *dernière étude*, 11 PQB et 11 sujets contrôles ont effectué 4 expériences sans feedback visuel. Dans les deux premières, ils devaient soit détecter un mouvement de leur mandibule induit par un mécanisme placé à l'intérieur de leur bouche, soit en estimer l'amplitude (Daliri, Prokopenko, & Max, 2013). Dans la troisième, ils devaient réaliser un mouvement de la mandibule le plus petit possible. Dans la quatrième, ils devaient reproduire le plus précisément possible une amplitude d'ouverture de la bouche préalablement induite par le mécanisme placé à l'intérieur de leur bouche. Aucune différence statistiquement significative n'a été observée entre les groupes. Cependant, certaines des PQB étaient moins précises que le sujet contrôle le moins précis aux expériences 3 et 4. Cela suggère que lorsque certaines PQB ont à leur disposition uniquement l'information proprioceptive pour contrôler leur mouvement et qu'elles doivent réaliser un mouvement minimal de leur mandibule ou maintenir en mémoire la position de leur mandibule afin de la rouvrir à une amplitude donnée, elles réalisent des mouvements plus amples et commettent plus d'erreurs.

2.1.3 Les études portant sur la perturbation du feedback proprioceptif

Dans deux dernières études, les capacités proprioceptives des PQB ont été testés par le biais d'une tâche dans laquelle il leur était demandé de réaliser un mouvement de la

mandibule le plus précis possible (de 18mm) avec ou sans perturbation de l'information proprioceptive. La perturbation de l'information proprioceptive était induite via un vibrateur placé sur le masséter gauche (muscle situé au niveau des joues et qui s'étend lors de l'ouverture de la bouche) des participants. Une telle vibration perturbe la source principale d'information proprioceptive de la mandibule (les fuseaux neuromusculaires) et donne l'illusion d'un allongement excessif du muscle et, donc, de déplacements de la mandibule plus amples qu'ils ne le sont. Il est donc attendu que cette méthode induise chez les participants une tendance à exécuter des mouvements de trop petite amplitude. En faveur de l'hypothèse selon laquelle les PQB utilisent moins efficacement les informations proprioceptives, les PQB étaient moins influencées par la vibration que les sujets contrôles dans une première étude (Loucks & De Nil, 2006a). Cependant, ces résultats n'ont pas été répliqués dans une étude proposée quelques années plus tard (Loucks & De Nil, 2012).

2.2 Les données neurofonctionelles

Ces 30 dernières années, le développement des techniques de neuro-imagerie nous a permis d'accumuler de nombreuses données concernant d'éventuelles anomalies cérébrales associées au bégaiement. Il parait donc intéressant de vérifier si les PQB ont des anomalies anatomiques ou fonctionnelles dans les parties du cerveau supposées jouer un rôle important dans la proprioception orofaciale et/ou dans l'utilisation de cette information pour le contrôle moteur de la parole.

Avant de discuter ces résultats, il semble utile d'introduire quelles sont les aires cérébrales supposées impliquées dans la proprioception. Cette question étant encore débattue à l'heure actuelle, et ne pouvant pas en faire une revue exhaustive dans ce chapitre, nous nous référerons aux propositions théoriques de Guenther et collaborateurs (2016). Depuis plusieurs années, ces chercheurs ont développé un modèle neuro-computationnel du contrôle moteur de la parole (Direction Into Velocities of the Articulators : DIVA) capable de simuler la production et le développement de la parole. Ce modèle est biologiquement plausible, c'est-à-dire qu'il intègre des régions cérébrales supposées impliquées dans la génération des mouvements articulatoires de la parole. Dans ce modèle, le gyrus supramarginal et le gyrus postcentral ventral jouent un rôle important dans l'utilisation et l'intégration des informations proprioceptives pour le contrôle moteur de la parole.

Plusieurs données permettent de penser que les personnes qui bégaient ont des anomalies dans ces aires cérébrales. Une méta-analyse de 17 études d'imagerie cérébrale (Belyk et al., 2015), en particulier, a mis en évidence que les épisodes de bégaiement des PQB sont accompagnés d'une diminution de l'activation du gyrus supramarginal, en comparaison à leur parole fluente. Une autre étude menée par Brown (2015) utilisant une méthode par imagerie spectroscopique proche infrarouge fonctionnelle (ISPIf) a permis de mettre en évidence une diminution d'activation au sein du gyrus supramarginal chez les participants qui bégaient dans une tâche de dénomination alors qu'une augmentation était observée chez les participants contrôles. Enfin, Sahin et al. (2005) rapportent le cas de deux patients présentant un bégaiement acquis à la suite d'un accident vasculaire cérébral affectant sélectivement l'intégrité du gyrus supramarginal chez un patient et l'intégrité du gyrus supramarginal et du gyrus angulaire chez l'autre.

3. Discussion

L'information proprioceptive a un rôle important dans le contrôle moteur de la parole. En effet, grâce à cette information, le locuteur peut vérifier rapidement si le son qu'il va émettre correspond bien à ses intentions en se basant sur la position de ses articulateurs. Nous savons également qu'au cours du développement normal le contrôle de la parole passe progressivement d'un contrôle auditif à un contrôle proprioceptif.

Dans la littérature sur le bégaiement, de nombreuses données suggèrent que les difficultés de parole rencontrées par les personnes qui bégaient seraient dues, en partie, à une forte dépendance au feedback auditif (Max et al., 2004 ; Tourville et al., 2008). Cette dépendance excessive pourrait s'expliquer par un trouble proprioceptif. Les quelques études comportementales et neurofonctionnelles qui se sont intéressées à cette question ont en effet mis en lumière des difficultés proprioceptives ainsi que des anomalies anatamo-fonctionnelles du système proprioceptif chez au moins certaines personnes qui bégaient. Plusieurs études ont mis en évidence qu'une partie des PQB sont en difficulté non seulement lorsqu'elles doivent identifier une forme géométrique placée dans leur cavité buccale, mais aussi lorsqu'elles doivent réaliser des mouvements précis avec leurs articulateurs (mandibule, langue et lèves) (Archibald, & De Nil, 1999 ; Daliri, et al., 2013 ; De Nil, & Abbs, 1991, Loucks, & De Nil, 2006b, 2012 ; Loucks, De Nil, & Sasisekaran, 2007 ; Martin et al., 1981 ;

Stewart et al., 1985). Par ailleurs, lorsque certaines PQB doivent réaliser des mouvements précis de la mandibule, en disposant d'informations proprioceptives perturbées induite par un vibrateur placé sur le masséter gauche, leur performance se dégrade moins que celles des sujets contrôles (Loucks, & De Nil, 2006a). Cela suggère que les PQB utilisent moins efficacement l'information proprioceptive

Ces données sont étayées par quelques études neurofonctionnelles qui ont mis en évidence des anomalies anatomo-fonctionnelles au niveau du gyrus supramarginal, zone supposée jouer un rôle important dans l'utilisation et l'intégration des informations proprioceptives pour le contrôle moteur de la parole selon le modèle DIVA (Belyk et al., 2015 ; Brown, 2015 ; Guenther, 2016 ; Sahin et al., 2005).

Cependant, ces difficultés ne semblent pas caractériser l'ensemble de la population des PQB. En effet, certaines études n'ont pas été en mesure de mettre en évidence de différence significative entre les groupes alors qu'elles utilisaient des paradigmes semblables (Daliri, et al., 2013 ; Loucks, & De Nil, 2006a, 2012 ; Stewart et al., 1985) et, dans toutes les études, certaines des PQB obtenaient des performances comparables à celles des sujets contrôles. Ces données suggèrent donc qu'il est plausible que la dépendance excessive à l'information auditive observée chez les PQB soit la conséquence d'un trouble proprioceptif. Elles ne permettent pas, néanmoins, d'établir une relation de cause à effet certaine.

Par ailleurs, bien que certaines données ont montré des difficultés proprioceptives chez certaines PQB, nous manquons encore de précisions quant à l'origine et l'étendue de ces difficultés (i.e. troubles proprioceptifs généraux ou propres à certains articulateurs ? Les difficultés proprioceptives observées sont-elles dues à une différence de bas niveau, comme un nombre moins important de récepteurs perceptifs comme les fuseaux neuromusculaires, ou à une différence dans la capacité des PQB à utiliser l'information proprioceptive pour le contrôle moteur de la parole ?) ainsi que sur le rôle joué par ces difficultés proprioceptives dans la genèse du bégaiement (i.e. Les troubles proprioceptifs sont-ils la cause ou la conséquence du bégaiement ? Est-ce parce que les PQB parlent moins qu'elles ont une moins bonne proprioception ?).

Ainsi, mieux comprendre les mécanismes à l'origines des troubles proprioceptifs observés chez certaines PQB et relier ces difficultés aux disfluences typiques du bégaiement,

via une analyse complète de leurs profils cognitif, moteur et langagier, permettrait d'améliorer l'efficacité des prises en charge.

Chapter 3. Is stuttering confined to articulated speech?

Abstract

This study explored whether adults who stutter (AWS) show phonological working memory difficulties independently of speech motor execution or broader working memory weakness. Using spoken, written, and visuospatial working memory tasks, group and individual analyses were conducted across twenty AWS and forty matched controls. Results revealed that, as a group, AWS tend to experience phonological working memory difficulties within general working memory impairment. However, distinct patterns were observed at the individual level. These findings highlight the importance of individual-level analyses and broad cognitive assessment for understanding the possible heterogeneous mechanisms underlying stuttering.

1. Introduction

Stuttering is probably one of the most studied speech disorders. However, despite the significant contributions of numerous researchers and the amount of data collected over the past century, our understanding of the mechanisms underlying stuttering still remains limited.

Although stuttering is commonly described as a fluency disorder that primarily affects speech motor control, some researchers propose that people who stutter (PWS) may experience difficulties extending beyond overt motor processes. Notably, several studies point to weaknesses in phonological working memory. According to Baddeley (2003), phonological working memory is supported by the phonological loop, a specialized system for the temporary storage and manipulation of speech-based information. The phonological loop comprises a phonological store, which briefly retains speech information for about two seconds, and an articulatory rehearsal process, which actively refreshes these memory traces either covertly (in the mind) or overtly (through spoken repetition).

Potential difficulties in these phonological working memory processes among PWS have been primarily assessed through nonword repetition tasks in which participants listen to a spoken nonword stimulus and are then asked to repeat it exactly as they heard it. To accomplish this, individuals must temporarily store the unfamiliar sequence in the phonological store while continuously refreshing it through articulatory rehearsal. This process enables the construction and maintenance of a novel phonological and articulatory plan until the nonword can be accurately articulated. These tasks have been administered in both children who stutter (CWS) and adults who stutter (AWS). Twelve studies in CWS have produced mixed results: while eight reported significantly poorer nonword repetition performance among CWS compared to controls (Anderson & Wagovich, 2010; Anderson, Wagovich, Hall, 2006; Hakim & Ratner, 2004; Oyoun et al., 2010; Pelczarski & Yaruss, 2016; Spencer & Weber-Fox, 2014; Sugathan & Maruthy, 2019; see also Ofoe et al., 2018 for a review), four found no significant group differences (Bakhtiar et al., 2007; Sasisekaran & Byrd, 2013; Smith et al., 2012; Vahab et al., 2014). For AWS, three of five studies found poorer performance on nonword repetition tasks compared to typically fluent controls, mainly for longer nonwords (six or seven syllables; Byrd et al., 2012; Byrd, McGill, & Usler, 2015; Sasisekaran & Weisberg, 2014), while two studies using shorter nonwords (one to four

syllables) found no group differences (Smith et al., 2010; Sasisekaran, 2013; see also Sasisekaran et al., 2013).

Overall, these studies suggest that PWS (as a group) may show difficulties on nonword memory tasks and thus that PWS might exhibit phonological working memory difficulties. However, the evidence remains inconsistent, with several studies reporting no significant group differences. Although nonword repetition tasks are widely used to assess phonological working memory in PWS, a number of methodological factors complicate the interpretation of these findings. Four key considerations are outlined below that should be taken into account when interpreting these results.

First, nonword repetition tasks also involve the production of overt speech. As a result, it is difficult to determine whether poorer performance in PWS arises from difficulties during the memory (or maintenance) stage, the execution stage, or both. To address this limitation, only one study (Byrd et al., 2015) attempted to isolate the two stages by including both vocal and non-vocal nonword memory tasks. The vocal task corresponded to a typical nonword repetition task, whereas in the non-vocal task, participants first heard a target nonword, followed by a set of three nonwords presented auditorily. They were then asked to manually select the nonword that matched the target by pressing the computer key 1, 2, or 3. Byrd and collaborators (2015) reported that AWS showed significantly reduced accuracy compared to controls in the vocal condition at longer syllable lengths (7 syllables), with a moderate effect size (partial $\eta^2 = .244$). In contrast, no group difference was observed in the non-vocal condition, and the effect size was minimal (partial $\eta^2 = .032$). Although this pattern is suggestive, performance in the non-vocal condition was near ceiling for both groups, suggesting that the task may not have been sensitive enough to detect subtle differences in phonological memory. Without a more demanding non-vocal task, or a direct interaction test across vocal and non-vocal conditions, it remains unclear whether group differences are truly restricted to phonological working memory.

Second, a nonword repetition task typically requires participants to listen to the spoken nonword and then repeat it exactly as they heard it. However, there is evidence that PWS may have difficulties accurately perceiving speech sound properties (see Foundas et al., 2004; Koca et al., 2024; Prestes et al., 2017; Robb, Lynn, & O'Beirne, 2013; Shao, Bakhtiar, & Zhang, 2022). This raises the possibility that poor performance on nonword

repetition tasks may reflect limitations in perceiving the acoustic features of the stimulus, rather than, or in addition to, deficits in working memory.

Third, research has shown that PWS tend to perform significantly worse than controls across various cognitive domains (Yairi, 2007). Thus, it is possible that the difficulties observed in phonological working memory reflect more general memory impairments.

Fourth, although group-level approaches have often indicated a trend toward phonological working memory difficulties in PWS, results remain inconsistent, with some studies finding no significant differences between stutterers and non-stutterers. One possible explanation for this inconsistency is that comparing mean group performance may obscure individual differences; not all PWS may experience difficulties with nonword repetition tasks. Indeed, it is widely recognized that PWS display heterogeneous profiles across the various cognitive, linguistic, and motor abilities required for fluent speech. Therefore, group-level analyses commonly used in previous research cannot determine whether all PWS in a given study exhibit the observed difficulties. To date, individual data are only available in the study by Spencer and Weber-Fox (2014), which presented the performance of CWS with persistent stuttering and those who recovered. According to the authors, their findings suggest that difficulties in phonological working memory may serve as a predictive factor for the persistence of stuttering. However, individual data revealed that only 37% of the CWS with persistent stuttering scored below the performance range of the recovered group, and only on tasks involving longer nonword sequences. This underscores the importance of considering individual variability in future research.

Importantly, addressing these methodological limitations is critical not only for clarifying the source of observed difficulties, but also for informing broader theoretical debates about the nature of stuttering. Demonstrating that PWS may exhibit difficulties in the covert processes of phonological working memory—processes that precede and are independent from overt articulation—would challenge the view that stuttering is solely a motor-based disorder. Instead, such evidence would suggest that cognitive-linguistic factors could play a role in the disorder's underlying mechanisms.

To address the methodological limitations discussed above, the present study proposes a nonword memory task with two response modalities: written (non-vocal) and spoken

(vocal), to better distinguish between phonological working memory and motor execution processes (Point 1). To minimize the potential confound of speech perception difficulties, the nonword stimuli will be presented in written form rather than auditorily (Point 2). To account for the possible influence of broader cognitive difficulties (Point 3), a visuospatial memory task is included as a control condition. Finally, to account for the heterogeneity within the cognitive, linguistic, and motor abilities that underly the fluent speech of PWS, this study assesses both group-level differences and individual profiles among PWS.

The design allows for testing competing predictions. If (some) AWS perform poorer across all three tasks to a similar extent, this would suggest a broader, domain-general memory difficulty (Eichorn, Hall, & Marton, 2023). If they (some) perform poorer in both phonological memory conditions (written and spoken) relative to the visuospatial task, this pattern would point toward a phonological working memory impairment. In contrast, if performance is specifically lower in the spoken condition compared to the written and visuospatial conditions, this would suggest that difficulties arise primarily during speech production. Some AWS may, however, perform well across all three tasks, implying that stuttering may arise from additional impairments not captured by the abilities assessed in the present study.

2. Methods

2.1 Participants

Twenty adult who stutter (AWS; 13 men and 7 women, aged 20-58 years, $M_{age} = 31$, $SD = 10.5$; years of education 2-8, $M = 4.65$, $SD = 2.3$) and forty control participants (20 men and 20 women, aged 20-60 years, $M_{age} = 28.2$, $SD = 9.1$; years of education 1-8, $M = 4.3$, $SD = 1.7$) took part in the study (see Table 1 for further information).

All AWS included in the study met the following criteria: they were diagnosed with stuttering by a speech therapist in accordance with the recommendations of the Belgian health authorities (AWS 3–8, 12–14, 16, 17, 20) or the French health authorities (AWS 1, 2, 9–11, 15, 18, 19). All participants self-identified as people who stutter (PWS) and reported that stuttering impacts their daily lives. None reported neurological events indicative of a neurogenic origin for their stuttering.

As outlined in Table 1, both the age at which the diagnosis of stuttering was made and the severity of stuttering varied among participants. With the exception of AWS 12, all had received some form of speech-language therapy. Except for AWS 6 and 15, all participants reported using various strategies—such as slowing their speech, incorporating frequent pauses, and using fillers—to improve fluency and reduce stuttering at the time of testing. In some cases, stuttering was not formally diagnosed until adolescence. These participants reported experiencing speech disfluencies during childhood, but these did not interfere with their lives until social, academic, or communicative demands increased during adolescence or adulthood. For some participants, stuttering was not detectable using the SSI-4 at the time of testing. The low SSI-4 scores observed for these individuals may be attributed to ongoing therapeutic intervention, as speech therapy is known to reduce the frequency, duration, and physical concomitants of stuttering, thereby directly impacting SSI-4 total scores. There is a well-established consensus in the literature regarding the high intra-individual variability of stuttering severity among PWS. For example, Constantino et al. (2016) demonstrated that stuttering frequency and SSI-4 severity scores can fluctuate considerably from day to day, sometimes dropping below the SSI-4 threshold of 10 used to indicate the presence of stuttering. Similar findings were reported by Frankford et al. (2022) who also included participants with stuttering severity scores below 10 (see also Tichenor et al., 2017).

Participants' developmental, neuropsychological, neurological, psychiatric, along with their family history of stuttering, were assessed through comprehensive interviews. Additionally, AWS provided information on the age at which stuttering was diagnosed, their history of therapy, and the strategies they employed in daily life and during testing to enhance fluency and minimize stuttering occurrences, if any.

All participants in the study were native French speakers with normal or corrected-to-normal vision and normal hearing. None reported significant developmental delays, learning disorders such as dyslexia, developmental language disorders, or dyspraxia, or attentional disorders. Among the control participants, there were no reports of neurological or psychiatric conditions, nor any family history of stuttering. Within the group of AWS, three individuals reported some neurological history: AWS 10 had a history of restless legs syndrome, AWS 11 and 17 experienced tics during adolescence, and AWS 13 was undergoing treatment for epilepsy at the time of testing. Five AWS participants reported a

psychiatric history: AWS 7, 8, 13, and 16 had experienced depression, with AWS 13 continuing treatment during the study, and AWS 1 was being treated for generalized anxiety disorder. Additionally, eleven AWS participants reported a familial history of stuttering, with affected relatives spanning first-degree (AWS 4, 6, 7, 14), second-degree (AWS 1, 2, 4, 5, 14, 15), third-degree (AWS 8), and fourth-degree (AWS 9, 20) family members. The neurological and psychiatric conditions reported by some AWS, such as tics, restless legs syndrome, or past depression, were not expected to affect performance on the study's tasks, while ongoing treatment for epilepsy or anxiety was anticipated to exert, at most, broad and non-specific effects across the measures. This was not considered problematic, as the experimental design allowed each participant's performance on phonological working memory tasks to be directly compared to their own results on a matched visuospatial control task, thereby minimizing the risk that general cognitive difficulties might be mistaken for specific deficits in speech-related domains. Importantly, because the study focused on individual cognitive profiles rather than group-level statistics, no single participant, including those with comorbid conditions, could influence the overall findings. All data are reported transparently, enabling readers to interpret each case independently and in line with a Bayesian approach to scientific reasoning.

Table 1

Participants information

Participants	Age	Sex	Years of education[1]	% SS reading[2]	% SS speech[3]	Duration score (s)[4]	Physical concomitants[5]	SSI4[6]	Stuttering severity[7]	Age of detection[8]
AWS 1	20	F	3	0.5	1.1	0.7	5	13	Very mild	18
AWS 2	30	F	8	0.7	1.2	0.76	3	11	Very mild	3
AWS 3	28	M	5	5.4	2.9	1.3	3	18	Mild	6
AWS 4	58	M	3	0.2	0.7	1.1	2	10	Very mild	3
AWS 5	53	M	8	4	5.8	2.4	6	25	Moderate	3
AWS 6	34	M	8	0.2	6	2	7	19	Mild	3
AWS 7	47	M	3	0.52	1.35	2	7	18	Mild	2,5
AWS 8	30	F	5	3.7	2.4	2.9	4	20	Mild	10
AWS 9	25	F	5	0.3	0.2	0.6	1	5	< Very mild	4
AWS 10	24	F	5	0.4	2.8	2.3	7	19	Mild	4
AWS11	22	M	2	0.5	1.6	0.2	2	9	< very mild	3
AWS12	29	M	5	0.3	2	0.2	1	6	< very mild	8
AWS 13	35	M	3	0.6	0.3	0.6	4	10	Very mild	3
AWS14	33	M	3	0.2	2.4	0.3	2	7	< very mild	3
AWS 15	27	M	2	0.3	0.4	0.2	5	7	< Very mild	3
AWS 16	20	F	2	2.2	5.8	1.93	6	22	Mild-moderate	9
AWS 17	27	M	8	0	4.4	1.3	7	18	Mild	13
AWS 18	33	F	8	0.2	1.2	0.2	1	6	< very mild	8
AWS 19	22	M	2	0.8	2.8	0.6	4	13	Very mild	16
AWS 20	24	M	5	7.4	11	3.4	13	36	Severe-very servere	7

Note. [1]Years of education: total years completed from after high school. [2-4]Measures extracted from participants' speech samples (for detail, see Evaluation of stuttering severity, below). [2] Percentage of stuttered syllables in the reading sample. [3] Percentage of stuttered syllables in the spontaneous speech sample. [4] Average duration of the three longest stuttering events in the spontaneous speech sample. [5] Severity of physical concomitants. [5-6] Stuttering severity score and category according to the SSI4 convention. [7] Reported age at which the term "stuttering" was first applied to the individual's speech difficulties, whether by the individual, a caregiver, a teacher, or a professional, in years.

2.2 General procedure

The study was approved by the biomedical ethics committee of the Cliniques universitaires Saint-Luc, Brussels, Belgium (#B403201734030). All participants provided written informed consent prior to participation. Each participant completed two individual sessions, each lasting approximately 45 minutes, during which they were administered the

same battery of tasks assessing verbal and visuo-spatial memory. Testing was conducted individually in a quiet environment, either at the participant's home, in the laboratory, in a speech therapist's office, or online. Twenty control participants and AWS 3, 4, 5, 6, 7, 8, 11, 12, 13, 14, 16, 17, 18, and 20 were tested by the first author, while the remaining twenty control participants were tested by an undergraduate speech and language therapy student. AWS 1, 2, 8, 9, 10, and 15 were tested by two experienced speech and language therapists. All computerized experiments were administered using PsychoPy software (versions 1.84.2 and 2024 1.5) (Peirce, 2009) or Pavlovia (https://pavlovia.org/).

2.2.1 Evaluation of stuttering severity

The evaluation of stuttering severity involved the collection of two audio- and video-recorded speech-language samples. For the first sample, participants were invited to read aloud "as if you were reading to a child" a passage of the book "*Le Petit Prince*" (Saint-Exupéry, 2002 pp. 21-23; 567 syllables) presented on a computer screen. For the second sample, participants were asked to answer open-ended questions on various topics such as their professional occupation, future perspectives and, when necessary, hobbies and vacation. The aim was to collect at least 5 minutes of conversational speech.

These speech samples were analyzed as follows. First, orthographic transcripts of the samples were generated following a slightly modified version of the VALIBEL transcription convention (Bachy et al., 2007). The changes were the following: (1) when the agreement of the past participle in the feminine was expected but not heard, the agreement form was reported, (2) pauses, fillers (e.g., /œ/, /œm/), paraverbal elements (e.g., coughs, laughs, etc), and repetitions of sounds and syllables were recorded as well as their duration. The transcriptions were then carefully reviewed by three master's students from the speech therapy program at UCLouvain, who had received training from the first author specifically for this task. Any discrepancies in the transcriptions were discussed and resolved through consensus.

On this basis, and following the SSI-4 (Riley, 2009) convention, we counted the number of "expected" syllables in each transcript—that is, the number of syllables the transcript would have contained if produced fluently. In other words, stuttered repetitions of a syllable were counted as a single syllable. We then created a coding system to identify 12

disfluency types from the audio and video-recordings: five stuttering-like disfluencies (SLDs; blocks, prolongations, monosyllabic word and syllable repetition if there are more than 1 repetitions and/or 1 single repetition accompanied with clear tension, phoneme repetition, and clusters), two strategies (avoidance/escape behaviors), and six other types of disfluency (hesitations, pauses and fillers, interjections, phrase repetition, monosyllabic word and syllable repetition if there is only one repetition and there is no tension, multisyllabic word repetition, revisions). Then, we measured the duration of all the SLDs (in Praat; Boersma & Weenink, 2024). For each participant we measured (1) the percentage of SLDs in the reading sample; (2) the percentage of SLDs in the conversational speech sample; (3) the average duration of the three longest SLDs of the spontaneous speech sample; (4) the severity of physical concomitants, rated according to the SSI4 convention. Finally, these measures were converted to a severity score and category according to the SSI4 convention (see Table 1).

2.2.2 Stimuli

Phonological working memory

A nonword memorization task was used to assess participants' phonological memory. Sixteen five-syllable and sixteen six-syllable sequences (n = 32) were selected for the task. Within each syllable-length group, half of the nonwords began with a vowel and half with a consonant. All nonwords were manually constructed to follow the phonotactic rules of French, respecting typical syllabic structures and avoiding unusual consonant clusters or phoneme combinations that would be implausible or unpronounceable in French. While no formal lexical verification was conducted, care was taken during construction to avoid sequences resembling real words or frequent morphemes. Most of the syllables (81.8%) followed a CV (consonant–vowel) structure, which is the most common syllabic pattern in French. Additionally, 9.1% of the syllables followed a V (vowel-only) structure, and 9.1% followed a CVC (consonant–vowel–consonant) structure. Each nonword was presented in white Arial font on a grey background (hex code: #808080), using PsychoPy's 'norm' units. The font size was set so that the stimulus height was 0.1 relative to the screen height, meaning it occupied approximately 10% of the vertical display.

Visuospatial working memory

A computerized task based on the Corsi Block-Tapping Test (Corsi, 1972) was used to assess participants' visuospatial working memory. Stimuli were a set of thirty-two ensembles of nine black

squares randomly positioned on a medium grey background (hex code: #808080) in such a way that they did not overlap. Each square measured approximately 1.5 degrees of visual angle in both height and width.

2.2.3 Tasks

Phonological working memory

The nonword memorization task included two response conditions: a spoken and a written response condition. Each condition began with a short familiarization phase consisting of two practice trials. The structure of each trial was as follows: (1) a nonword was displayed at the center of the screen for 6 seconds, during which participants were asked to memorize it silently, without engaging in any articulatory movements. It was explicitly emphasized that the memorization should take place internally. The experimenter monitored participants during the training phase to ensure they followed this instruction; (2) after the 6 seconds, the nonword disappeared and a fixation cross appeared at the center of the screen for 12 seconds, during which participants were instructed to continue retaining the nonword silently; (3) after the retention period, the fixation cross disappeared and a message prompted participants to either write down (written condition) or pronounce aloud (spoken condition) as many syllables as they could remember. In the nonvocal condition, participants used capital letters and wrote on a sheet of paper placed in front of them. In the vocal condition, participants' responses were audio recorded using a Dictaphone for later analysis. The five- and six-syllable nonwords were intermixed throughout the task. Participants managed the progression from one trial to the next themselves, with no time limit imposed between trials.

Visuospatial working memory

Each trial of the visuospatial task followed the same structure: (1) the nine black squares were displayed simultaneously for 2 seconds, allowing participants to encode their spatial positions; (2) a sequence of five or six squares then lit up in red, one at a time for one second. Participants were instructed to memorize the order in which these locations were illuminated. After the sequence, participants were prompted to reproduce the order by clicking on the corresponding squares using a computer mouse.

When a square was clicked, it briefly flashed to confirm the input. Clicking elsewhere on the grey background produced no feedback, ensuring that only valid responses were acknowledged. The

five- and six-square sequences were separated into two blocks, with the five-square sequences presented first, followed by a short pause, and then the six-square sequences. A 3 second countdown preceded each block. As in the phonological task, participants progressed through the trials at their own pace, with no time limit between trials.

2.2.4 Procedure

The nonword memorization and visuospatial memory tasks were administered in two separate testing sessions, each lasting approximately 45 minutes. Session 1 included collection of anamnestic information, language samples and administration of the written condition of the nonword memorization task. Session 2 included the spoken condition of the nonword memorization task, followed by the visuospatial memory task. All participants completed the sessions in the same order, except for AWS 3 who performed the spoken condition during the first session and the written condition in the second one due to technical difficulties.

All tasks were presented on a screen monitor placed approximately 57 cm from the participant. AWS 1, 2, 4, 6, 7, 8, 9, 10, 12, 13, 15, 16, 17, 19, 20 were tested in person. Additionally, to increase the sample size and include individuals unable to travel to the testing site, we conducted part of the study remotely via online sessions.

Both the nonword memorization task and the visuospatial task were administered in person and online. For in-person testing, the nonword memorization task was controlled using PsychoPy version 1.84.2 (Peirce, 2009), and the visuospatial task was run using PsychoPy version 2024.1.5 (Peirce, 2009) on a Windows 10 laptop.

For online testing, the nonword memorization task was administered by the experimenter via screen sharing, while the visuospatial task was hosted and controlled through the Pavlovia platform (https://pavlovia.org/). For the online version of the visuospatial task, we included a screen calibration step to ensure consistent display size across participants' devices. Specifically, we used the "screen scale" routine available through the PsychoPy community (see: https://osf.io/8fhqk/), which allowed participants to calibrate their screen by adjusting a reference object (e.g., a credit card or bank card) to match its actual size on screen. This step ensured that the spatial dimensions of the stimuli were appropriately scaled, making the visuospatial task comparable across devices. Participants completed the tasks on their own computers under the real-time supervision of the

experimenter through a video conferencing platform. Aside from the mode of delivery, the procedure remained identical across testing formats.

2.2.5 Scoring

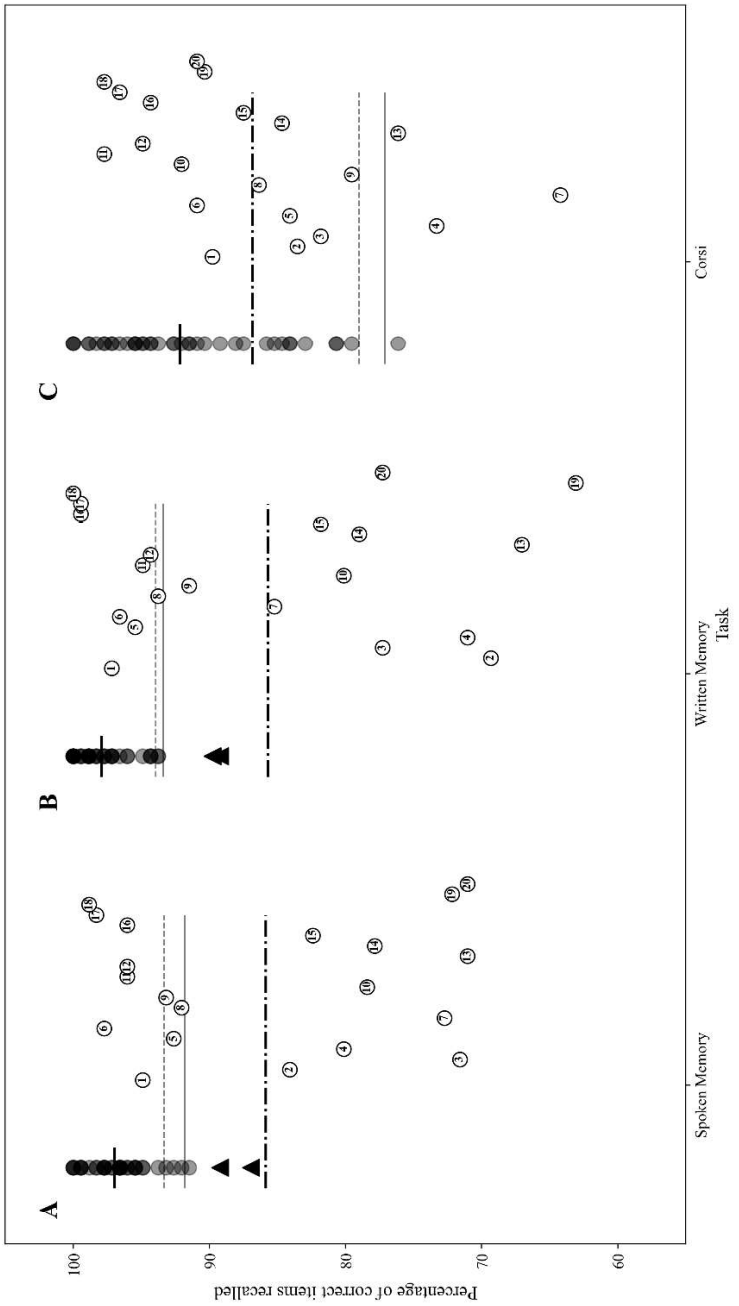
Accuracy for each trial was calculated as the number of syllables (in the phonological tasks) or squares (in the visuospatial task) correctly recalled in the correct serial position, relative to the total number of items in the target sequence. The percentage of correct items recalled was then computed.

3. Analysis and results

The first step in data analysis was to examine if some control participants exhibited performance levels that appeared abnormally high or low. While the exclusion of "outliers" is a common practice in the field, this process is somewhat arbitrary. Here, we opted for a systematic exclusion of outliers based on the 1.5 x interquartile range (IQR) rule, i.e., performances falling outside a range between 1.5 times the IQR below the first quartile and 1.5 times the IQR above the 3rd quartile. As a result, three control participants were excluded. The results of all participants are presented in Figure 1.

Figure 1

Results of all participants in the three tasks



Note. Percentage of correct items recalled for controls (grey filled circles) and AWS (white circles labeled with their ID) is shown for each condition of the memory tasks: the spoken (panel A) and written (panel B) conditions of the phonological working memory task, as well as the visuospatial task (Corsi; panel C). Grey lines indicate statistical thresholds, corresponding to the first (plain) and either to the fifth (spoken) or second (Written and Corsi) percentile (dashed) estimated percentiles of the control distribution. The mean performance of the control group and the AWS are indicated by a black solid line and a black dash-dot line, respectively.

In a first series of analysis, the sample of AWS successfully recruited enabled us to examine whether previously reported findings could be replicated, specifically whether AWS, as a group, demonstrated difficulties regarding the spoken condition of the phonological working memory task. This was assessed by comparing the mean percentage of correctly produced syllables in the AWS group to that observed in normo-fluent control participants. As a preliminary step, we investigated whether gender influenced the number of syllables correctly produced, given that gender differences have been reported in visuospatial memory performance (Orsini et al., 1986). This analysis was essential to determine whether data from both genders could be pooled or should be analyzed separately. An independent samples *t*-test was conducted using IBM SPSS Statistics (version 25) to assess whether the mean percentage of correctly produced syllables differed between male and female control participants. All model assumptions were checked and met. As no significant difference was found between genders, $t(36) = 0.19$, $p = 0.85$, data from both genders were subsequently pooled in the group-level analyses. We then proceeded to address our main question and conducted an independent sample *t*-test to determine whether AWS, as a group, performed more poorly than controls on the spoken condition of the phonological working memory task (see Figure 1, panel A). The independent sample *t*-test revealed that AWS scored significantly lower than controls, $t(19,9) = 4,54$, $p < 0.001$. These results demonstrate that, as a group, AWS exhibit significantly reduced performance on spoken memory tasks compared to their typically fluent peers. This pattern aligns with previous research documenting similar group differences in phonological working memory performance.

However, the significant difference observed between PWS and controls on the spoken condition of the phonological working memory task does not rule out the possibility that the difficulties among AWS may arise from factors other than phonological working memory anomalies, such as mechanism required during speech production or broader working memory weaknesses. To clarify this, we first conducted two Mann–Whitney U tests to compare performance between AWS and controls on the written condition of the phonological working memory task and on the Corsi visuospatial memory task, given that the control group's distribution deviated from normality. These analyses revealed that AWS performed significantly more poorly than controls on both tasks: written memory, $U = 124$, $Z = -4.21$, $p < .001$; and Corsi, $U = 259$, $Z = -2.2$, $p = 0.03$. Then, to determine whether AWS experienced greater difficulties in any specific task, an analysis of variance (ANOVA) using GROUP (AWNS, AWS) as a between-subjects factor and TASK (spoken, written, corsi) as a within-subjects factor was conducted. Because the control group's distribution deviated from normality for the written phonological memory condition and the Corsi task, we first applied an Aligned Rank Transformation (ART; Leys & Schumann, 2010) to the data prior to the ANOVAs. The analysis revealed a significant main effect of GROUP, $F(1, 55) = 25.9$, $p < 0.001$, $\eta^2 = 0.32$ and TASK, $F(1,4) = 5.2$, $p = 0.016$, $\eta^2 = 0.086$. These results indicate that AWS recalled a significantly lower percentage of items across tasks compared to controls, and that the percentage of items recalled varied significantly between the different memory tasks for both groups. More interestingly, there was a trend toward a significant interaction between GROUP and TASK, $F(1.32, 72.81) = 3.49$, $p = 0.054$, $\eta^2 = 0.06$, suggesting that the magnitude of group differences varied as a function of task type.

To further explore this trend, post hoc comparisons were conducted by performing three separate 2×2 factorial ANOVAs, each with GROUP (AWNS, AWS) as the between-subjects factor and a pair of TASKS (spoken vs. written, spoken vs. Corsi, written vs. Corsi) as the within-subjects factor. For spoken memory vs. Corsi, there was a significant main effect of GROUP, $F(1, 55) = 27.6$, $p < 0.001$, $\eta^2 = 0.334$, but neither the main effect of TASk, $F(1, 55) = 3.3$, $p = 0.07$, $\eta^2 = 0.057$, nor the INTERACTION, $F(1, 55) = 3.0$, $p = 0.088$, $\eta^2 = 0.052$, reached significance. For written memory vs. Corsi, significant main effects were found for GROUP, $F(1, 55) = 26$, $p < 0.001$, $\eta^2 = 0.324$, and TASK, $F(1, 55) = 7.6$, $p = 0.008$, $\eta^2 = 0.121$ as well as a significant INTERACTION, $F(1, 55) = 4.6$, $p = 0.036$, $\eta^2 = 0.077$. For spoken vs. written memory, there was a significant main effect of GROUP, $F(1, 55) = 14.6$, p

< 0.001, $\eta^2 = 0.21$ but neither the main effect of TASK, $F(1, 55) = 3.2$, $p = 0.08$, $\eta^2 = 0.055$, nor the INTERACTION, $F(1, 55) = 0.6$, $p = 0.44$, $\eta^2 = 0.011$, was significant.

In sum, the group level analyses indicate that although AWS, as a group, exhibit broad working memory difficulties—demonstrated by robust group effects across all three tasks—they also show particular weaknesses in phonological working memory that are independent of response modality (i.e., not solely attributable to speech production demands). Specifically, AWS performed significantly worse than controls on the written condition of the phonological working memory task compared to the Corsi task, with a similar trend toward weaker performance on the spoken condition versus the visuospatial task. More importantly, comparison of the two conditions of the phonological working memory task demonstrates that AWS consistently performed worse than controls across both tasks.

In a second series of analyses, we sought to determine whether the pattern observed in the group level analysis was similar for all the AWS taken individually. To this end, we applied the Crawford and Howell (1998) modified *t*-test. Given that the control group's distribution deviated from normality and exhibited moderate skewness and slight leptokurtosis for the written condition (skewness = -0.89; kurtosis = 0.76) and the Corsi task (skewness = -0.79; kurtosis = - 0.03), we followed the recommendations of Crawford et al. (2006) and used a more stringent significance threshold ($p < .02$) to limit the inflation of the Type I error rate associated with non-normal control data. Crawford et al. (2006) demonstrated that, under conditions of moderate skewness and leptokurtosis and with control group sizes similar to ours, the actual Type I error rate using the Crawford and Howell method at a 2% significance level falls between 3.59% and 3.78%.

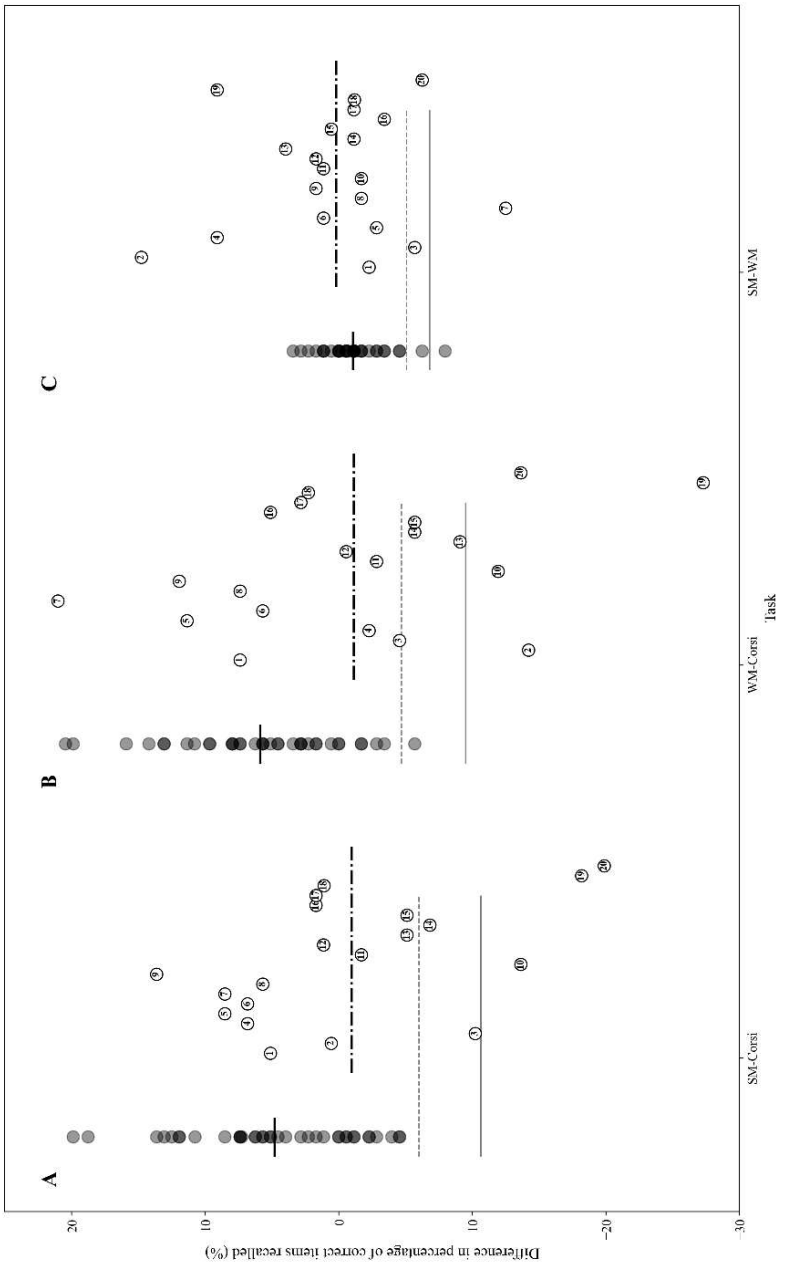
As for the group level analysis, to further explore whether AWS, individually, experienced greater difficulties in any specific task, a series of modified *t*-tests on the difference scores between tasks were conducted. The individual differences between tasks are displayed in Figure 2, and the results of the modified *t*-tests, along with their respective p-values, are presented in Table 2. The objective of these analyses was to determine whether the performance of AWS in the spoken condition was distinct from their performance in the written and visuospatial (Corsi) conditions, and crucially, whether these differences were atypical relative to those observed in the control group. By comparing each AWS's difference scores (e.g., Spoken–Corsi, Written-Corsi, Spoken–Written,) to the distribution of

corresponding difference scores among controls, we aimed to identify individual cases where the discrepancy between tasks exceeded what would be expected in the typically fluent population. The results reveal heterogeneity in working memory profiles among AWS. Several AWS (AWS 1, 6, 11, 12, 16, 17, 18) demonstrated no difficulties across any of the assessed tasks, performing within the range, and even above average for some (AWS 16, 17 and 18), observed for the control group. In contrast, a subset of participants (AWS 2, 3, 4, 5, 7, 9, 13) exhibited broad deficits, with impaired performance evident across all working memory measures. Some AWS displayed specific difficulties restricted to phonological working memory tasks (AWS 8, 10, 14, 15, 19, 20), with one individual (AWS 20) showing particularly pronounced impairment when required to produce spoken responses, suggesting a motor component to the difficulty. Furthermore, among those with broad working memory difficulties, the impairments tended to be more severe for most of them in the phonological working memory conditions (AWS 2, 3, 4, 7, 13), and for certain individuals (AWS 3 and 7), the deficit was especially marked when the task involved overt speech production.

These findings underscore the diversity of profiles within the AWS group, highlighting that while some individuals exhibit generalized or modality-specific working memory impairments, others perform comparably to typically fluent controls. Again, this pattern emphasizes the importance of considering individual cognitive diversity when interpreting group-level differences in working memory among AWS.

Figure 2

Individual differences between tasks



Note. The difference scores in the percentage of correct items recalled for controls (grey filled circles) and AWS (white circles labeled with their ID) is shown for every task comparison: Spoken - Corsi (panel A), Written - Corsi (panel B) and Spoken - Written (panel C). Grey lines indicate statistical thresholds, corresponding to the first (plain) and to the fifth percentile (dashed) estimated percentiles of the control distribution. The mean performance of the control group and the AWS are indicated by a black solid line and a black dash-dot line, respectively.

Table 2

Modified t-test results and p-values for each AWS on phonological working memory and visuospatial working memory tasks and their difference scores

AWS	SM t (*p*)	WM t (*p*)	Corsi t (*p*)	SM-Corsi t (*p*)	SM-WM t (*p*)	WM-Corsi t (*p*)
1	-0.97 (.17)	-0.46 (.33)	-0.39 (.35)	0.05 (.48)	0.5 (.31)	.24 (.41)
2	**-6.03 (<.001)****	**-14.98 (<.001)****	-1.4 (.09)	-0.66 (.26)	**6.69 (<.001)****	**-3.2 (.001)****
3	**-11.89 (<.001)****	**-10.83 (<.001)****	-1.68 (.05)	**-2.35 (.01)****	**-1.94 (.03)***	**-1.66 (.05)***
4	**-7.89(<.001)****	**-14.09 (<.001)****	**-3.06 (<.002)****	0.32 (.38)	**4.29 (<.001)****	-1.3 (.1)
5	**-2.04 (.02)***	-1.35 (.09)	-1.31 (.1)	0.59 (.28)	-0.74 (.23)	0.88 (.19)
6	0.36 (.36)	-0.75 (.23)	-0.2 (.42)	0.32 (.38)	0.94 (.18)	-0.03 (.49)
7	**-11.36 (<.001)****	**-6.68 (<.001)****	**-4.53 (<.001)****	0.59 (.28)	**-4.82 (<.001)****	**2.42 (.01)****
8	**-2.3 (.01)****	**-2.23 (.02)***	-0.94 (.18)	0.14 (.45)	-0.26 (.4)	0.24 (.41)
9	**-1.77 (.04)***	**-3.42 (<.001)****	**-2.04 (.02)***	1.39 (.09)	1.18 (.12)	0.97 (0.17)
10	**-8.69 (<.001)****	**-9.35 (<.001)****	-0.02 (.49)	**-2.89 (.003)****	-0.26 (.4)	**-2.84 (<.001)****
11	-0.44 (.33)	-1.64 (.06)	0.9 (.19)	-1.02 (.16)	0.94 (.18)	-1.39 (.09)
12	0.44 (.33)	-1.94 (.03)	0.44 (.3)	-0.57 (.29)	1.18 (.12)	-1.03 (.16)
13	**-12.15 (<.001)****	**-16.16 (<.001)****	**-2.6 (.01)****	-1.55 (.07)	**2.14 (.02)***	**-2.39 (.01)****
14	**-8.96 (<.001)****	**-9.94 (<.001)****	-1.22 (.12)	**-1.82 (.04)***	-0.02 (.49)	**-1.84 (.04)***
15	**-6.83 (<.001)****	**-8.46 (<.001)****	-0.76 (.23)	-1.55 (.07)	0.7 (.25)	**-1.84 (.04)***
16	0.44 (.33)	0.73 (.24)	0.35 (.37)	-048 (.32)	-0.98 (.17)	-0.12 (.45)
17	0.63 (.27)	0.73 (.24)	0.72 (.24)	-048 (.32)	-0.02 (.49)	-0.48 (0.32)
18	0.88 (.19)	1.03 (.16)	0.9 (.19)	-.058 (.28)	-0.04 (.49)	-0.57 (.29)
19	**-11.62 (<.001)****	**-18.24 (<.001)****	-0.3 (.38)	**-3.6 (<.001)****	**4.29 (<.001)****	**-5.29 (<.001)****
20	**-12.15 (<.001)****	**-10.83 (<.001)****	-0.2 (.42)	**-3.86 (<.001)****	**-2.18 (.02)***	**-3.11 (.002)****

Note. Results of modified *t*-tests and associated *p*-values for each adult who stutters (AWS) on the spoken, written, and visuospatial (Corsi) working memory tasks, and on their difference scores. The analyses aimed to determine whether individual AWS exhibited significant deficits in any task and whether their performance differences between tasks (e.g., Spoken–Corsi, Written–Corsi, Spoken–Written) were atypical relative to the control group distribution. Asterisks indicate statistically significant *p*-values. For the Written Memory and Corsi tasks, $p < .02$ is marked with *, and $p < .01$ with **. For the Spoken Memory condition and for all task difference scores, $p < .05$ is marked with *, and $p < .01$ with **. AWS = adults who stutter; SM = Spoken Memory; WM = Written Memory.

4. Discussion

This study investigated whether adults who stutter (AWS) exhibit specific impairments in phonological working memory that are independent of speech motor execution processes, as well as broader working memory difficulties. To address this, both group- and single-case-level analyses were conducted, allowing examination of overall group patterns and individual performance profiles. This study has led to two interesting findings.

First, AWS as a group tend to experience weaknesses in phonological working memory, a finding consistent with previous research utilizing nonword repetition tasks (Anderson & Wagovich, 2010; Anderson et al., 2006; Byrd et al., 2012, 2015; Hakim & Ratner, 2004; Oyoun et al., 2010; Pelczarski & Yaruss, 2016; Sasisekaran & Weisberg, 2014; Spencer & Weber-Fox, 2014; Sugathan & Maruthy, 2019; see also Ofoe et al., 2018 for a review). However, these phonological working memory difficulties were observed within the context of broader working memory impairment.

Second, distinct profiles were observed among AWS. (1) In line with the group-level results, several AWS (2, 3, 4, 7, 13) showed broad working memory weaknesses, with more pronounced impairments in the phonological working memory conditions; (2) other AWS (1, 6, 11, 12, 16, 17, 18) performed within the normative range across all tasks; (3) a few AWS (5, 9) also exhibited broad working memory weaknesses, without these being more pronounced in any specific condition; (4) some AWS (8, 10, 14, 15, 19, 20) showed selective deficits restricted to phonological working memory tasks; (5) finally, for certain AWS (3, 7, 20), their weaknesses were particularly pronounced when the task required overt speech production.

Overall, our findings have important implications. First, the identification of five distinct working memory performance profiles among AWS carries important theoretical implications. For some individuals, broad general working memory weaknesses suggest that such difficulties could play a role in the emergence of stuttering, given that fluent speech depends on the temporary retention and efficient retrieval, processing, and storage of linguistic, sensory, and motor information (Eichorn, Hall, & Marton, 2023). In addition, the finding that certain AWS show pronounced difficulties with the storage and maintenance of phonological information—even in the absence of overt speech—challenges models that

attribute stuttering primarily at the level of speech motor execution. Rather, these observations indicate that deficits at pre-articulatory stages, particularly those involving the retrieval, storage, and maintenance of phonological information, may play a more significant role in stuttering, at least for some individuals. Conversely, the presence of AWS who performed within the normative range across all tasks suggests that, for these speakers, stuttering may originate from difficulties at earlier stages of speech planning, rather than from deficits in working memory.

Second, these findings highlight the need to incorporate individual-level analyses in future research, as AWS demonstrate diverse patterns of working memory performance, it would not be surprising to observe similar variability across other domains essential for fluent speech. Such an approach is critical for identifying potentially heterogeneous loci of dysfunction within the cognitive mechanisms underlying speech production in PWS. However, larger-scale studies and intervention research are necessary to determine whether any of these patterns reflect causal relationships with stuttering or merely represent comorbidities or compensatory mechanisms.

5. Conclusion

This study examined whether AWS show specific phonological working memory deficits independent of speech motor execution, or broader working memory weaknesses. Using spoken, written, and visuospatial working memory tasks, the results revealed that AWS, as a group, tend to exhibit phonological working memory difficulties within a broader working memory weakness. However, individual analyses showed substantial variability, with some AWS performing within the typical range and others displaying either broad or selective deficits.

These findings suggest that stuttering may not arise from a single underlying deficit but from multiple potential loci of dysfunction within speech-related cognitive mechanisms. Notably, phonological working memory difficulties observed in the absence of speech output challenge purely motor accounts of stuttering. Future studies should integrate individual-level analyses and larger, longitudinal designs to determine whether these cognitive patterns contribute causally to stuttering or reflect either comorbidities or compensatory mechanisms. A broader evaluation of the various cognitive domains required for fluent speech at the

individual level could also further improve our understanding of the potentially heterogeneous cognitive bases of stuttering.

Chapter 4. Testing theories of stuttering: cognitive heterogeneity but shared orosensory tactile weaknesses

Abstract

The cognitive underpinnings of stuttering are likely heterogeneous. However, there is very little data speaking directly to this heterogeneity in the current literature. Virtually all studies have focused on just one or a few hypotheses, and most of them have only analyzed group differences. To contribute to identifying the (potentially heterogeneous) cause(s) of stuttering, 12 adults who stutter (AWS) were administered a battery of tests assessing an extensive set of linguistic, auditory, somatosensory, rhythmic and motor abilities known to underpin fluent speech. Using a case study approach, the performance of each AWS was compared against that of 20 control participants to establish individual cognitive profiles. The AWS has heterogeneous cognitive profiles. Yet, surprisingly, all exhibited some form(s) of orosensory tactile weakness, and for some this was the only weakness observed. This finding highlights the value of cognitive profiling in stuttering and suggests a stronger association between stuttering and orosensory tactile disorders than previously assumed.

Reference

Van Thorre, A., Patri, J. F., & Vannuscorps, G. (2025). Testing theories of stuttering: Cognitive heterogeneity but shared orosensory tactile weaknesses. *Cognitive Neuropsychology*, 1–40. https://doi.org/10.1080/02643294.2025

1. Introduction

Stuttering is a complex disorder characterized by considerable interindividual variability in symptoms (Arndt & Healey, 2001; Blood, 1985; Riley & Riley, 1980; Riley & Riley, 2000; Yaruss, La Salle, & Conture, 1998), brain function and anatomy (e.g., Foundas et al., 2004; Kell et al., 2009), and treatment responses (e.g., Alm, 2004; Maguire et al., 2004). This variability has led to growing speculation that, across individuals, stuttering may arise from different types of dysfunction(s) within the functional architecture supporting fluent speech—that is, the constellation of perceptual, sensorimotor, linguistic, cognitive and motor components that together enable speech production (Alm, 2004; Civier et al., 2013; Max et al., 2004; Yairi, 1990, 2007).

Yet, this growing recognition has so far had little impact on the way research is conducted in the field. Stuttering remains predominantly studied using group comparisons, where the average performance of a group of people who stutter (PWS) on one or a few tasks is compared to that of typically fluent individuals (Seery et al., 2007; SheikhBahaei et al., 2023; Yairi, 2007). This approach poses significant challenges, potentially obscuring the diverse underlying causes of stuttering (e.g., McCloskey & Caramazza, 1988; Shallice & Evans, 1978; SheikhBahaei et al., 2023). As noted over sixty years ago by St. Onge and Calvert (1964) 'if stuttering is not a unitary phenomenon, then the investigator is almost certainly working with a spurious sample, resulting in a nonrandom mix of two or more undefined populations', leading to inconsistent and contradictory findings across studies—an ongoing source of confusion in this field (for discussion see also Seery et al., 2007; Yairi, 2007).

More importantly, the group-comparison approach seems inadequate to inform the theories and hypotheses that it intends to test. If the goal is to determine whether a specific process (e.g., speed of semantic access) is a causal contributor of stuttering, the critical issue is not whether *on average* PWS perform in the same range than normo-fluent controls in a task assessing that process. If the hypothesis is that the process *x* is a cause of stuttering in all the PWS, the question is whether it functions significantly less effectively in all the PWS (than in normo-fluent controls). And if the hypothesis is that process *x* may be a cause of stuttering in at least some stutterers, then, the question is whether it functions significantly less effectively in at least some PWS (than in normo-fluent controls). The group-comparison approach is ill-suited to answer these questions. Group studies may

reveal group differences even if only a few, or none, of the PWS show individually significant deficits (see for instance, De Nil & Abbs, 1991;see also results below). In contrast, studies may fail to reach statistical significance while some PWS exhibit clear difficulties (e.g., Daliriet al., 2013; Stager, 1990). In sum, causal factors with low prevalence will be difficult, if not impossible, to detect, and the pattern detected at a group level may not describe well any single member of the group – an issue classically referred to as the "averaging artefact" (Shallice & Evans, 1978). Answering these types of questions requires a case-by-case analysis.

Group-comparison studies' reliance on a small number of tasks additionally implies that there is very little information in the current literature on how PWS perform across tasks and domains. Research has identified group differences between PWS and normo-fluent controls in various domains, including in psycholinguistic, auditory, somatosensory, timing, and motor tasks (see below). However, interpreting these findings remains challenging. One possibility is that most PWS experience difficulties across all or most of these domains, potentially due to domain-general factors like attentional deficits or performance anxiety, or to a widespread effect of genetic and neurological factors underlying stuttering. These difficulties could also stem from the impact of stuttering on cognitive development. Another possibility, however, is that different PWS struggle in different (combinations of) domains, suggesting the existence of cognitive subtypes of stuttering (e.g., linguistic, auditory, somatosensory, timing, or motor subtypes). Still another possibility is that stuttering is indeed causally homogeneous. The same core deficit would then be observed in all the PWS (i.e., it is necessary) and it could be the only deficit found in at least some PWS (i.e., it is sufficient). In this case, heterogeneity would emerge from different collections of associated disorders. Teasing apart these alternatives requires having an overview of the strength and weaknesses of each PWS across the different potentially important domains of skills and abilities, and then of the cognitive heterogeneity and similarities between participants.

This study aimed to take an initial step in this direction by employing a multiple single-case approach. A key strength of this approach, which has provided valuable insights into cognitive heterogeneity in some other neurodevelopmental disorders such as dyslexia (Ramus et al., 2003; White et al., 2006), dyscalculia (Träff et al., 2017) and autism (Towgood et al., 2009), is its focus on in-depth analysis of individual cases, rather than relying on pooled data across participants. In particular, we will examine the unique patterns of strengths and weaknesses in 12 individuals who stutter across five key domains of abilities, each known to

support speech production and hypothesized to contribute to stuttering (see below for details). Collectively, these individual patterns of strengths and weaknesses will be referred to hereafter as each individual's *cognitive profile*. In this context, *cognitive profile* refers to an inventory of strengths and weaknesses that provides a window into the integrity of the functional architecture supporting fluent speech—that is, the constellation of perceptual, sensorimotor, linguistic, cognitive and motor components that together enable speech production. These cognitive profiles will then be analyzed to identify or rule out potential causal factors underlying each participant's stuttering. Finally, the cognitive profiles of all PWS will be analyzed collectively to evaluate the role of various factors and theoretical models. Special attention will be given to determining whether any specific factor could be considered necessary and/or sufficient to explain stuttering in these individuals. By doing so, we hope to begin charting the landscape of heterogeneous strengths and weaknesses in PWS, thereby providing a foundation for future research into whether stuttering is underpinned by homogeneous causal factors or by distinct cognitive subtypes. Moving forward, progress will depend on establishing which of the weaknesses identified in the cognitive profiles of PWS play a causal role in stuttering, and whether different causal pathways may be identified. Addressing these questions will require larger samples and experimental designs that go beyond the descriptive approach taken in the present study.

1.1 Psycholinguistic abilities

Speech production proceeds through the successive access to multiple independent levels of representation (Caramazza, 1997; Levelt et al., 1999; Shelton & Caramazza, 1999). Object picture naming, for instance, entails an initial stage of perceptual processing and object recognition, followed by semantic access, lexical selection, retrieval of syntactic-grammatical information, encoding of metrical and segmental phonological information, and phonetic encoding. Semantic access involves retrieving stored knowledge about the object. Then, the object's name needs to be retrieved from the mental lexicon. While the exact dynamic of this "lexical selection" remains debated (Mahon et al., 2007) this lexical selection is often assumed to result from a competitive process by which the speaker selects the most appropriate word by inhibiting lexical competitors (e.g., Shao, Roelofs, & Meyer, 2012; Shao, Meyer, & Roelofs, 2013; Shao et al., 2014). Lexical selection provides access to grammatical-gender information (e.g., Navarrete et al., 2006). The next step is that of the

phonological encoding of the segmental and metrical properties of the word. For example, for the word [puppy], segmental information involves retrieving the four phonemes that compose it (/p/ /ʌ/ /p/ /i/) and metrical information specifies its prosodic frame—for instance, that it is disyllabic with stress on the first syllable (in English). This is followed by a stage of "phonetic encoding", where all the phonemes of the word are related to their corresponding classes of actions, like "lip closing" for /p/. Finally, this motor plan is executed, and the classes of actions are translated in precisely timed muscle contractions of the vocal articulators.

A disorder of any of these stages has the potential to influence speech production. Accordingly, stuttering has long been hypothesized to emerge from a range of different psycholinguistic disorders (see Seery et al., 2007, for review), ranging from abnormally slow semantic processes (Bosshardt, 2006; Bosshardt & Fransen, 1996) to lexical selection difficulties (Newman & Ratner, 2007; Pellowski & Conture, 2005), abnormally weak inhibitory control of lexical selection (Maxfield, 2020, 2021), abnormally slow phonological encoding (Wingate, 1988), abnormally frequent phonological errors (Kolk & Postma, 1997; Postma & Kolk, 1993; Sasisekaran et al., 2006; Weber-Fox, 2001; Weber-Fox et al., 2004), abnormally inefficient phonetic encoding (Howell, Au-Yeung, & Sackin, 2000) or speech motor skills (Caruso, Abbs, & Gracco, 1988; Kleinow & Smith, 2000; Max, Caruso, & Gracco, 2003; Van Lieshout, Hulstijn, & Peters, 1996). Although there are various ways to relate each of these disorders to stuttering (Maxfield, 2020, for instance), according to a popular hypothesis abnormally slow processing at any of these stages may lead to stuttering if, when the execution of the previous syllable or word is complete, due to upstream slowing the motor plan for the next segment is not yet available (Howell, 2007, 2010).

1.2 Auditory skills

Auditory processing is crucial for speech production, as the accurate perception of sound properties—such as frequency, intensity, duration, and acoustic features like formant configuration—enables effective auditory control of speech (Guenther, 2016; Hickok, Houde, & Rong, 2011). Accordingly, auditory processing deficits have long been hypothesized to contribute to stuttering (Bloodstein, 1995; Guitar, 1998; Koca, Belgin, & Ölçek, 2024; Lawrence & Barclay, 1998; Smith, 1990; Webster & Lubker, 1968). One way such deficits might contribute

is by interfering with the development of stable internal models that predict the auditory outcomes of speech motor commands. These deficits would also, by definition, compromise the processing of auditory feedback during speech. Together, unstable internal models and hampered auditory processing could undermine fluency by increasing the likelihood of perceived mismatches between expected and actual auditory feedback—even when no true error has occurred—thereby triggering unnecessary corrective commands that disrupt smooth speech (Chang et al., 2016; Lotfi et al., 2020; Max et al., 2004; Wieland et al., 2015).

In line with this hypothesis, differences between PWS and normofluent controls have been reported in a large variety of central auditory tasks such as in tone detection (Hampton & Weber-Fox, 2008), pitch pattern sequence test, dichotic listening, speech in noise (Asal & Abdou, 2014; Curry & Gregory, 1969; Hall & Jerger, 1978), prediction during speech comprehension (Gastaldon et al.; 2024a, b), backward masking, duration pattern and gap in noise tests (Howell et al., 2000; Koca et al., 2024; Lotfi et al., 2020). Moreover, PWS have been reported to show weaker than normal compensatory responses to experimental modifications of the auditory feedback (Bradshaw et al., 2021; Cai et al., 2012; Cai et al., 2014). The possible role of auditory factors in stuttering is also supported by the fact that fluency in PWS is often improved when their speech is masked (Sutton & Chase, 1961; Wingate, 1970), shifted in frequency (Howell, El-Yaniv, & Powell, 1987) or delayed (Armson et al., 2006; Kalinowski et al., 1993; Wingate, 1970).

1.3 Somatosensory skills

The domain of somatosensory skills encompasses two distinct faculties: proprioception, which pertains to the perception of one's articulatory position and movement, and touch, which involves the detection and processing of tactile stimuli. Together, these faculties contribute to speech planning and control (Abbs & Gracco, 1983, 1984; Folkins & Abbs, 1975; Gracco & Abbs, 1989; Kent, 2024). Producing the phoneme /t/, for instance, requires contact between the anterior part of the tongue and the alveolar ridge. Planning a motor command to achieve this position requires an accurate estimate of the initial position of one’s tongue and jaw, which may be derived from a combination of proprioceptive and tactile information. And then, the detection of a contact with the desired force at the desired locations on the tongue and palate may be monitored to ensure that the speech goal will be successfully achieved, as well as to trigger the release of the subsequent speech unit (Civier et al., 2013).

Given its role in speech planning, control and sequencing, it is not surprising that it has for long been suggested that stuttering could be rooted in a deficient orosensory processing system (Mysak, 1960). Van Riper (1971), for instance, suggested that stuttering may emerge in children who fail to transfer from auditory to somatosensory control of speech (see also Lane & Tranel, 1971). Although the available empirical data in this area remains limited (and at times contradictory), some findings have emerged regarding possible tactile and proprioceptive processing deficits in PWS. In the tactile domain, although some studies have shown that PWS tend to exhibit typical spatial acuity measured by two-point discrimination tests (Jensen et al., 1975; Zsilavecz, 1981), others have reported a slightly higher vibrotactile detection threshold on the tongue (Fucci et al., 1985, 1991; Petrosino et al., 1987). And in the proprioceptive domain, several studies have reported that PWS make significantly larger jaw, tongue and lower lip movements – but not finger movements – than control participants during minimal movement tasks (Archibald & De Nil, 1999; De Nil & Abbs, 1991), that their movements are less precise (more variable) when asked to open their jaw to a specific amplitude (Loucks, De Nil, & Sasisekaran, 2007; Loucks & De Nil, 2006b), that their orofacial motor skills are less affected by the perturbation of proprioceptive information induced by tendon vibration (Loucks & De Nil, 2006a; but see Loucks & De Nil, 2012), and that they performed less accurately than controls in oral stereognosis (Class, 1956; Gunderson, 1972; Martin et al., 1981; Stewart, Evan, & Fitch, 1985; Zsilavecz, 1981; but see Carpenter, & Sommers; 1987; Cook et al., 2011; Hutchinson, & Ringel, 1975; Jensen et al., 1975).

1.4 Rhythmic and timing skills

Speech production is widely assumed to rely on the ability to generate internal rhythmic signals across various time scales (e.g., sub-syllabic, syllabic, phrase) (Allen, 1973; Barbosa, 2007; Byrd & Saltzman, 2003; Cummins, 2009; Cummins & Port, 1998; Dilley, Wallace, & Heffner, 2012; Etchell, Johnson, & Sowman, 2014; Jones, 1976; Large & Jones, 1999; Martin, 1972; McAuley & Jones, 2003; Port, 2003; Schwartze & Kotz, 2013).

Accordingly, a rhythmic deficit has long been suggested to contribute to stuttering (Alm, 2004; Andrews et al., 1983; Cooper & Allen, 1977; Garnett et al., 2023; Hulstijn et al., 1992; Kent, 1984; Max & Yudman, 2003; Olander, Smith, & Zelaznik, 2010; Wieland et al., 2015;

Zelaznik, Smith, & Franz, 1994). For instance, stuttering has been suggested to result from an inaccurate "neural clock" (Cooper & Allen, 1977), a reduced ability to generate temporal patterns for speech perception and production (Kent, 1984), an unreliable mechanism for timing control (Andrews et al., 1983) and, more recently, from a deficit in the ability to internally generate consistent rhythmic motor behaviors (Olander et al., 2010) or from a dysfunction within a brain network that supports internal timing, resulting in reliance on external timing cues to sequence speech movements (Etchell et al., 2014; see also Alm, 2004). In line with this hypothesis, several studies have reported that speech production in PWS is temporally more variable (e.g., Cooper & Allen, 1977; Kleinow & Smith, 2000; Max & Gracco, 2005; Smith & Kleinow, 2000) and that stuttering is associated with a deficient rhythmic and/or timing abilities in both verbal and non-verbal domains, and in both production and perception tasks such as in paced finger tapping and rhythm discrimination tasks (Archibald & De Nil, 1999; Borden, 1983; Cooper & Allen, 1977; Falk, Müller, & Bella, 2014; Hulstijn et al., 1992; Kleinow & Smith, 2000; Max & Gracco, 2005; Olander et al., 2010; Smith & Kleinow, 2000; Smits-Bandstra, De Nil, & Rochon, 2006; Westphal, 1933; Wieland et al., 2015; Zelaznik et al., 1997; but see Hilger, Zelaznick, & Smith, 2016; Brown et al., 1990; Max & Yudman, 2003; Neef et al., 2011; Slis et al., 2023; Zelaznik et al., 1994).

This hypothesis is also in keeping with the observation that PWS often become more fluent when synchronizing their speech to an external pacing signal, such as an isochronous metronome (Wohl, 1968), when speaking in unison with another person (''choral speech'') (Adams & Ramig, 1980; Ingham & Carroll, 1977), and when singing (Glover et al., 1996). That external predictable temporal cues can considerably aid PWS seems to be pointing at a possible causal link. Another line of research suggesting a possible causal link, comes from the finding that brain damage on the territory of the Internal Timing Network (i.e., the basal ganglia and SMA) may result in stuttering (e.g., Ciabarra et al., 2000; Tani & Sakai, 2011).

1.5 Motor execution

Once an abstract motor plan has been defined, it must be executed. This involves the precise timing and coordination of muscle contractions to produce the desired sounds. Any impairments or disruptions at this latest stage will affect a person's ability to speak fluently. Not surprisingly, it has been proposed that stuttering may arise from an impaired speech motor system (Kleinow & Smith,

2000; Packman, 2012; Smith, 1999; Smith & Kleinow, 2000) and/or limitations in (speech) motor skills (Namasivayam & Lieshout, 2011; Peters, Hulstijn, &Van Lieshout, 2000). To be clear, stuttering is not viewed as a motor disorder such as dysarthria or dyspraxia. The motor system is intact in principle, but its performance is less stable and less efficient, especially under increased task demands, leading to breakdowns in fluency. These limitations may only become evident when complexity increases, such as when high demands for movement accuracy and speed are placed on the system. In line with this proposal, delayed initiation and slower execution in the speech domain (Peters, et al., 2000) as well as in non-verbal oral and in manual tasks have been reported in PWS (e.g., Archibald & De Nil, 1999; Borden, 1983; Hulstijn et al., 1992; Smits-Bandstra et al., 2006), suggesting that these motor weaknesses are not necessarily confined to the domain of speech (Namasivayam & Lieshout, 2011).

2. Methods

2.1 Participants

We recruited as many adults who stutter (AWS) as was possible for our team through social media, networks and professional and patient associations in a three-year period (between 2019 and 2022). Recruitment for studies involving extensive cognitive evaluations is particularly challenging, as participants are required to commit to numerous sessions—about 10 hours in total in our case. Twelve AWS (six men and six women, aged 20-47 years, M_{age} = 28.42, SD = 7.65; years of education 2-8 years, M = 4.17, SD = 2.2) and 20 control participants (six men and 14 women, aged 22-30 years, M_{age} = 26.25, SD = 2.59; years of education 3-8, M = 4.35, SD = 1.27) took part in the study (see Table 1 for further information).

Participants' developmental, neuropsychological, neurological, psychiatric, and orofacial histories, along with their family history of stuttering, were assessed through comprehensive interviews. Additionally, AWS provided information on the age at which stuttering was diagnosed, their history of therapy, and the strategies they employed in daily life and during testing to enhance fluency and minimize stuttering occurrences, if any.

All participants were native French speakers with normal or corrected-to-normal vision and normal hearing, verified through pure-tone audiometry. None of the participants

reported a history of significant developmental delays, learning disorders (e.g., dyslexia, developmental language disorder, or dyspraxia), or attentional disorders. None of the control participants reported neurological or psychiatric conditions, atypical orofacial anatomy, or a family history of stuttering.

All the AWS included in the study met the following criteria: they had received a formal diagnosis of stuttering from a licensed speech-language therapist based on the (very similar) criteria of either the Belgian (AWS 3–5, 8, 10, 11) or French health authorities (AWS 1, 2, 6, 7, 9, 12); they reported ongoing stuttering-related challenges in daily life; they self-identified as PWS; and they had no history of neurological events that could plausibly suggest a neurogenic origin for their stuttering, nor any accompanying conditions that might have prevented to interpret their performance on the tasks used in this study (e.g., a severe attention disorder).

As detailed in Table 1, the age at which stuttering was detected and the current severity of symptoms varied across participants. In some cases, stuttering was not formally diagnosed until adolescence, but age of diagnosis should not be confused with age of onset, and late detection does not necessarily indicate a different underlying condition. Milder forms of stuttering are often overlooked or misattributed until social, academic, or communicative demands increase in adolescence or adulthood. In a few cases, participants' stuttering was not detectable using the SSI-4 at the time of testing. However, a low SSI-4 score at a single point in time should not be taken as evidence that an individual does not stutter—especially given the well-documented variability in symptom severity among PWS and the fact that all the AWS participants had received some form of speech-language therapy. With the exception of AWS 3 and 10, all reported actively using fluency-enhancing strategies—such as slowing their speech, inserting pauses, and using fillers—to reduce disfluency during testing (see Table S1). Given their self-identification and experience of stuttering in daily life, their ability to limit the number of stuttering-like dysfluencies under controlled conditions reflects the effectiveness of these techniques, not the absence of an underlying disorder.

Among the AWS, seven reported a familial history of stuttering, with affected relatives spanning first-degree (AWS 3, 4), second-degree (AWS 1, 2, 9), third-degree (AWS 5), and fourth degree (AWS 6) family members. Three reported some neurological history:

AWS 7 had a history of restless legs syndrome, AWS 10 experienced tics during adolescence, and AWS 8 was undergoing treatment for epilepsy at the time of testing. Five AWS participants reported some psychiatric history: AWS 4, 5, and 10 had a past history of depression, AWS 8 was undergoing treatment for depression at the time of the study, and AWS 1 was being treated for generalized anxiety disorder. Finally, two AWS participants had atypical orofacial anatomy: AWS 8 was born with a cleft palate and has dental implants for teeth numbers 11, 12, 21, and 22; AWS 2 uses a dental prosthesis. In line with our inclusion criteria, these conditions were either not expected to affect performance in the tasks used in this study (tics, restless leg syndromes, past depression), to exert only broad, non-specific effects across all tasks and measures, if any (e.g., treatment for epilepsy) or to potentially influence performance in only one or a few specific tasks (atypical orofacial anatomy). This was not problematic given the study's goal and analytical framework. As discussed in the "General Analyses" section below, most tasks were designed to allow performance of each participant to be compared between their results on orofacial/speech-related tasks and matched control tasks. This approach minimizes the risk that general cognitive difficulties might be mistaken for specific deficits in speech-related domains. In addition, and importantly, because our study focuses on individual cognitive profiles rather than group-level statistics, no single participant, including those with accompanying conditions, distorts the overall findings. All data are transparently reported, allowing readers to interpret each case independently, consistent with a Bayesian perspective on scientific reasoning. In this spirit, the performance of the two participants with atypical orofacial anatomy (AWS 8 and AWS 2) is specifically addressed in the Results section on somatosensory function, ensuring that their profiles can be appropriately contextualized.

Table 1

Participants information

Participants	Age	Sex	Years of education[1]	% SS reading[2]	% SS speech[3]	Duration score (s)[4]	Physical concomitants[5]	SSI4[6]	Stuttering severity[7]	Age of detection[8]
AWS 1	20	F	15	0.5	1.1	0.7	5	13	Very mild	18
AWS 2	30	F	20	0.7	1.2	0.76	3	11	Very mild	3
AWS 3	34	M	20	0.2	6	2	7	19	Mild	3
AWS 4	47	M	15	0.52	1.35	2	7	18	Mild	2,5
AWS 5	30	F	17	3.7	2.4	2.9	4	20	Mild	10
AWS 6	25	F	17	0.3	0.2	0.6	1	5	< Very mild	4
AWS 7	24	F	17	0.4	2.8	2.3	7	19	Mild	4
AWS 8	35	M	15	0.6	0.3	0.6	4	10	Very mild	3
AWS 9	27	M	14	0.3	0.4	0.2	5	7	< Very mild	3
AWS 10	20	F	14	2.2	5.8	1.93	6	22	Mild-moderate	9
AWS 11	27	M	20	0	4.4	1.3	7	18	Mild	13
AWS 12	22	M	14	0.8	2.8	0.6	4	13	Very mild	16

Note. [2-4] Measures extracted from participants' speech samples (for detail, see Evaluation of stuttering severity, below). [1]Years of education: total years completed from middle school through to their highest level of formal education.[2] Percentage of stuttered syllables in the reading sample. [3] Percentage of stuttered syllables in the spontaneous speech sample. [4] Average duration of the three longest stuttering events in the spontaneous speech sample. [5] Severity of physical concomitants. [6-7] Stuttering severity score and category according to the SSI4 convention. [8] Reported age at which the term 'stuttering' was first applied to the individual's speech difficulties, whether by the individual, a caregiver, teacher, or clinician.

2.2 General procedure

The study was approved by the biomedical ethics committee of the Cliniques universitaires Saint-Luc, Brussels, Belgium (# B403201734030), and all participants gave their written informed consent to the study. Data collection was carried out from October 2019 to October 2022. All participants took part in six sessions lasting between 60 and 120 minutes. They were administered the same battery of tasks evaluating their speech, as well as various types of linguistic, auditory, somatosensory, rhythmic and motor abilities (Table 2). Participants were tested individually in a quiet room either at home, at the laboratory, or in a speech therapist's office. Control participants and AWS 1-6 were tested by the first author. AWS 7-12 were tested subsequently by two experienced speech and language therapists who were blind to the profiles of AWS 1-6. All the computerized experiments were controlled by the Psychopy software (version 1.84.2; Peirce, 2009).

Table 2

Domains, tasks and measures used in the study

Domain	Task	Modality	Condition	Effector	Dependent variable
Psycholinguistic	Object recognition				IES
	Semantic				IES & TC (IES Level(i) – IES Level(i-1))
	Grammatical-gender decision				IES & TC (IES Level(i) – IES Level(i-1))
	Segmental-phonological				IES & TC (IES Level(i) – IES Level(i-1))
	Metric-phonological				IES & TC (IES Level(i) – IES Level(i-1))
	Phonetic				IES & TC (IES Level(i) – IES Level(i-1))
Inhibitory control	Flanker	Manual			CE (IES congruent – IES incongruent)
		Verbal			CE (IES congruent – IES incongruent)
Auditory	Speech-in-noise				% of correct repetitions
	Dichotic aptitude				% of correct repetitions and REA
	Binaural advantage				Detection threshold (dB)
	Pitch pattern sequence				% of correct repetitions
	Duration pattern sequence				% of correct repetitions
Somatosensory	Point pressure sensitivity			Right and left wrist	Point pressure sensitivity threshold (mN)
				Inferior lip	Point pressure sensitivity threshold (mN)
				Right and left angle	Point pressure sensitivity threshold (mN)
				Right and left apex	Point pressure sensitivity threshold (mN)
	Force discrimination			Right and left wrist	Force discrimination threshold (dB)
				Inferior lip	Force discrimination threshold (dB)
				Right and left angle	Force discrimination threshold (dB)
				Right and left apex	Force discrimination threshold (dB)
	Thickness discrimination			Hand, jaw, and tongue	Thickness discrimination threshold (mm)
	Length discrimination			Thumb and tongue	Length discrimination threshold (mm)
Rhythm/timing	Synchronization	Manual			Rhythmic deviation (Hz) & variability (SD, in s)
		Verbal			Rhythmic deviation (Hz) & variability (SD, in s)
	Continuation	Manual			Rhythmic deviation (Hz) & variability (SD, in s)
		Verbal			Rhythmic deviation (Hz) & variability (SD, in s)
Motor execution	Maximum speed	Manual	Simple	Right and left hand	Maximal rate (Hz) and variability (SD, in s)
			Alternate	Right and left hand	Maximal rate (Hz) and variability (SD, in s)
		Verbal	Simple	Syllable	Maximal rate (Hz) and variability (SD, in s)
			Alternate	Syllable	Maximal rate (Hz) and variability (SD, in s)
		Orofacial	Simple	Praxis	Maximal rate (Hz) and variability (SD, in s)
			Alternate	Praxis	Maximal rate (Hz) and variability (SD, in s)

Note. The five domains investigated in the study, along with their corresponding tasks, modality, condition, effector tested, and respective dependent variables. Abbreviations: mN = millinewtons; dB = decibels; mm = millimeters; s= seconds; IES = inverse efficiency score; SD = standard deviation; TC = transition cost; CE = congruency effect; % = percentage; REA = right ear advantage; Hz = hertz.

2.3 Tasks and measures

2.3.1 Evaluation of stuttering severity

The evaluation of stuttering severity involved the collection of two audio- and video-recorded speech-language samples. For the first sample, participants were invited to read aloud "as if you were reading to a child" a passage of the book "*Le Petit Prince*" (Saint-Exupéry, 2002 pp. 21-23; 567 syllables) presented on a computer screen. For the second sample, participants were asked to answer open-ended questions on various topics such as their professional occupation, future perspectives and, when necessary, hobbies and vacation. The aim was to collect at least 5 minutes of conversational speech.

These speech samples were analyzed as follows. First, orthographic transcripts of the samples were generated following a slightly modified version of the VALIBEL transcription convention (Bachy et al., 2007). The changes were the following: (1) when the agreement of the past participle in the feminine was expected but not heard, the agreement form was reported, (2) pauses, fillers (e.g., /œ/, /œm/), paraverbal elements (e.g., coughs, laughs, etc.), and repetitions of sounds and syllables were recorded as well as their duration. The transcriptions were then carefully reviewed by three master's students from the speech therapy program at UCLouvain, who had received training from the first author specifically for this task. Any discrepancies in the transcriptions were discussed and resolved through consensus.

On this basis, and following the SSI-4 (Riley, 2009) convention, we counted the number of "expected" syllables in each transcript—that is, the number of syllables the transcript would have contained if produced fluently. In other words, stuttered repetitions of a syllable were counted as a single syllable. The size of the language samples read aloud were

of 579.44 syllables on average (range 568 – 600) for the controls and 579.9 syllables on average for the AWS (range 559 – 606). For the conversational speech sample, size of the samples was of 1135.88 syllables on average (range 908 – 1519) for the controls and 839.4 syllables on average for the AWS (range 378 – 1298).

We then created a coding system to identify 12 disfluency types from the audio and video-recordings: five stuttering-like disfluencies (SLDs; blocks, prolongations, phoneme repetition, clusters, and monosyllabic word and syllable repetition when involving more than one repetitions or a single repetition accompanied by clear tension), two strategies (avoidance/escape behaviors), and six other types of disfluency (hesitations/pauses/fillers, interjections, phrase repetition, multisyllabic word repetition, revisions, and monosyllabic word and syllable repetition limited to a single repetition without tension).Then, we measured the duration of all the SLDs (in Praat; Boersma & Weenink, 2024). For each participant we measured (1) the percentage of SLDs in the reading sample; (2) the percentage of SLDs in the conversational speech sample; (3) the average duration of the three longest SLDs of the spontaneous speech sample; (4) the severity of physical concomitants, rated according to the SSI4 convention. Finally, these measures were converted to a severity score and category according to the SSI4 convention (see Table 1).

2.3.2 Linguistic assessment: efficiency of access to successive levels of representation

Stuttering has been suspected to emerge from various types of psycholinguistic dysfunctions (Howell, 2007, 2010; Seery et al., 2007). To test for the presence of such disorder(s) in our participants, we designed a battery that comprised six tasks assessing participant's ability to retrieve information pertaining to six progressive levels of representation involved in picture naming. In those tasks, participants were presented with a set of pictures of familiar objects and successively asked to decide if the stimuli correspond to real objects (Level 1; "**object recognition** task"), if they were larger than a soccer ball (Level 2; "**semantic** task"), if the name corresponding to the object was feminine (Level 3; in French - "**grammatical-gender** task"), if the name included the sound /a/ (Level 4; "**segmental-phonological** task"), if the name had more than one syllable (Level 5; "**metric-phonological** task"), and if their lips would touch one another if they had to pronounce the object's name (Level 6; "**phonetic** task"). The object recognition task (Level 1) does not

assess a psycholinguistic function per se. However, it is a necessary step in object picture naming. Its inclusion allowed us to measure putative interindividual differences in object recognition and/or in the time needed to plan and execute a motor response once a decision has been made.

Stimuli for task assessing Levels 2 to 6 were the same 40 line-drawings of familiar objects (281x196 pixels), obtained from Snodgrass & Vanderwart (1980) and selected such as that in all tasks there was an equal number of accurate "yes" and "no" responses. For the "object recognition" task (Level 1), stimuli included 20 of the line-drawings used in tasks assessing Levels 2 to 6 and 20 line-drawings of unreal objects (281x196 pixels) obtained from Barbarotto et al. (2002).

The procedure was the same in the six tasks. Participants were introduced with the task and asked to provide their judgment as quickly and as accurately as possible by pressing one of two computer keys with their left ("n" key for "non" – "no" in French) or right ("o" key for "yes" – "oui" in French) index finger. Then, the task started. In each trial, participants saw a fixation cross for 1 second, followed by the presentation of a stimulus depicted with a size of ≈ 5 degrees of visual angle on the centre of the screen. The picture remained on the screen until a response (i.e. a button press) was recorded, which also initiated the next trial. The order of the stimuli was randomized between tasks, but identical across participants. Each task was preceded by a short familiarisation phase using different but similar stimuli (14 trials for task 1 and seven trials for tasks 2 to 6).

Before the six tasks, participants were familiarized with the type of task and the response mode. They were presented with 40 words, one at a time, and asked to quickly and accurately decide whether each word represented a color (e.g., "purple") or not (e.g., "dog"). Additionally, after the first task ("object recognition"), participants were shown the 40 line-drawings stimuli, each displayed one by one along with its corresponding written name. This procedure followed standard practices in studies of cognitive processes involved in picture naming (Damian, Vigliocco, & Levelt, 2001; Navarrete et al., 2013; Roelofs, 2006).

Participants' responses and response latencies (RL) were recorded. RL analyses were conducted on correct trials only (95.9 % and 93.4 % of the trials across all tasks for the controls and AWS, respectively). Additionally, trials with RLs deviating by more than 2

standard deviations from each participant's mean RL for correct responses in each task were discarded (4.5% and 4.7% of the remaining trials across all tasks for the controls and AWS, respectively). From this, we computed each participant's inverse efficiency score for each task. The inverse efficiency score is computed for each participant by dividing the mean response latency by the proportion of correct responses in a given condition (thus, the higher the score the poorer the performance). This score allows combining the measures of accuracy and speed into a single measure of processing efficiency. Thereby, it also allows between-group comparisons unbiased by potential speed-accuracy tradeoffs (Townsend & Ashby, 1978, 1983).

The performance of the AWS in the six tasks will be interpreted in light of two measures: (1) their inverse efficiency score (IES) for each task and (2) their difference in efficiency score between each task and the task assessing the previous processing level (IES Level(i) – IES Level(i-1); the "transitioning cost"). The joint consideration of these two measures is motivated by their complementary interpretative value. Efficiency serves as a diagnostic indicator for assessing an AWS's ability to access various levels of linguistic representation. High efficiency in a task targeting a specific level (e.g., the grammatical-gender decision task) suggests that difficulties at that level, as well as at preceding levels, are unlikely contributors to stuttering. Conversely, abnormally low efficiency in such a task is compatible with the hypothesis that a deficit in accessing information at that level may play a role in stuttering. However, considering efficiency alone is not sufficient, as a deficit detected at any given level may be due either to a difficulty to access information at that particular level of representation, or to a variety of other factors involved in the task, including general task-related slowness, difficulties in object recognition or attentional disorder, for instance. As a result, two participants with similarly low efficiency scores might differ significantly in the underlying causes of their performance.

Transition costs, in contrast, offer a more direct assessment of the cognitive processes targeted by each task. By isolating the added difficulty associated with accessing specific levels of linguistic representation, transition costs provide a clearer estimation of the origin of participants' difficulties, independent of confounding factors. However, considering transitioning costs in isolation could also present issues, as it might penalize participants who perform exceptionally well in certain tasks (e.g., AWS 2's performance in Tasks L3 and L4,

as discussed below). This could lead to the incorrect conclusion that a participant with normal-range performance in one task (e.g., the semantic task) has a deficit (e.g., a semantic deficit) simply because they performed exceptionally well in the preceding task.

Here, an AWS will be deemed to have a specific difficulty in retrieving information at a particular level (e.g., semantic difficulty) only if: (1) they perform significantly below the controls' performance in the task assessing that level (e.g., the semantic task) and (2) they show a significantly larger cost than the controls' in transitioning to this level from the previous one (e.g., object recognition).

2.3.3 Linguistic assessment: inhibitory control

As stuttering has also been suggested to be associated with difficulty to resolve lexical competition (Maxfield, 2020, 2021), linguistic assessment also included a **verbal Flanker** and a **manual Flanker** task (Eriksen & Eriksen, 1974). The verbal modality was designed to test for the hypothesis that AWS may exhibit speech-related inhibitory control difficulties (Maxfield, 2020, 2021; Maxfield et al., 2016). The manual modality aimed to determine whether these difficulties, if any, are specific to lexical competition or more general (Eggers, De Nil, & Van den Bergh, 2010, 2013; see also Wiltshire et al., 2020).

During the experiments, participants were placed in front of a monitor at a viewing distance of approximatively 57 cm. A medium grey background was displayed on the computer screen (hex color code: #808080). Stimuli were four different horizontal arrays of five arrows, measuring each about 1 degree of visual angle, depicted in white ink, and spaced 1.4 degrees of visual angle apart. The central arrow was centred on the centre of the monitor. There were four different stimuli: two "congruent stimuli" in which all the arrows pointed in the same direction (< < < < < and > > > > >) and two "incongruent stimuli" in which the central arrow pointed in a different direction (> > < > > and < < > < <).

In each trial, participants first heard a brief sound (440 Hz) for 250 ms, followed by the appearance of one of the four stimuli after a 500 ms delay. Participants were instructed to identify the orientation of the central arrow as quickly and accurately as possible. In the manual modality, participants responded by pressing the left or right arrow key of an external keyboard with their dominant hand's most comfortable fingers. In the verbal modality, participants responded by saying "gauche" or "droite" (French for "left" and "right") into a

sensitive microphone connected to a voice key controlled by PsychoPy software. The stimuli remained visible until a response was recorded or for a maximum of 1500 ms. In the manual modality, the subsequent stimulus was presented after a variable interstimulus interval, ranging from 500 to 1500 ms, following a uniform random distribution. In the verbal modality, the experimenter managed the transition between trials. This allowed the experimenter to monitor the functioning of the voice key and to manually record participants' responses.

Participants performed two blocks of 60 trials for each modality of the task (240 trials in total), following an ABBA order (A = manual modality). The first and second blocks were preceded by a short familiarisation phase (24 trials). Each block started by a 3-seconds countdown.

Participants' responses and response latencies (RL) were recorded and analyzed. In the verbal modality of the task, the RLs corresponded to the post-stimulus onset latency of the subject's vocalization. RL analyses were conducted on correct trials only (98.83% and 98.56% of the trials across all tasks for the controls and AWS, respectively). Additionally, voice key failure (1.45% and 1.77% for the controls and AWS, respectively) and trials with RLs deviating by more than two standard deviations from each participant's mean RL for correct responses in each task (4% and 4.4% of the remaining trials across all tasks for the controls and AWS, respectively) were discarded. Based on this, we calculated the inverse efficiency score for each participant in each modality of the task. The congruency effect was determined by subtracting the efficiency score of congruent trials from that of incongruent trials. A larger congruency effect reflects inhibitory control difficulty (Eriksen & Eriksen, 1974).

2.3.4 Auditory abilities

The assessment of both peripheral and auditory temporal processes included pure tone audiometry (Madsen Micromate 304), and six other tasks covering various aspects of auditory processing that have been suggested to be atypical in (at least some) PWS (Kekade & Valame, 2014; Koca et al., 2024). All these tests were administered via headphones (Sony MDR-XB550AP).

Pure tone audiometry was included to determine participants' basic hearing ability across a range of frequencies. In this test, participants were seated with their back to the experimenter, presented with tones of different frequencies and intensities and instructed to press a button when they detected a tone. Tones were presented in the following order: 1 kHz, 2 kHz, 4 kHz, 8 kHz, 500 Hz and 250 Hz to the right ear and then the left ear. The experiment followed a 3-down 1-up staircase procedure. For each frequency, the initial intensity was set at 30 dB SPL. If a participant failed to detect the tone in the first trial, the intensity was increased by 5 dB SPL. After three successive correct responses at a given intensity, the tone intensity was reduced by 5 dB SPL for the next trial. Any incorrect response resulted in an increase in tone intensity by 5 dB SPL. The experiment continued for each frequency until the tone intensity was gradually lowered to the point where the participant could no longer hear it. The tone intensity was then raised until the participant provided three successive correct responses, indicating their hearing threshold.

Central auditory processing encompasses a range of abilities, such as binaural integration and separation, temporal processing, and auditory pattern recognition, which collectively contribute to effective speech motor control and have been suggested to contribute to stuttering (Fox et al., 2020; Kekade & Valame, 2014; Lotfi et al., 2020; Shetty, 2016). The central auditory processing tasks used in this study were subtests of a validated and standardized French language central auditory battery (Demanez et al., 2003) designed based on the American Speech-Language-Hearing Association Consensus Statement on central auditory processing and models (ASHA, 2005). In all these subtests, stimuli were presented at 70 dB SPL.

The first task assessed participants' ability to identify spoken words embedded in white noise. **Speech-in-noise** recognition relies heavily on auditory temporal processing (Shetty, 2016), a skill that previous studies suggest may be weaker on average in PWS (Foundas et al., 2004; Shao, Bakhtiar, & Zhang, 2022), potentially contributing to reduced speech-in-noise recognition (Asal & Abdou, 2014). Speech-in-noise performance was evaluated using a list of 30 monosyllabic words (Lafon lists). The stimuli were presented binaurally at 70 dB SPL in two conditions: (1) embedded in white noise at 70 dB SPL (SNR =1) and (2) in isolation, without noise. Participants were instructed to repeat each word, and the accuracy of each specific phoneme within each word was verified by the experimenter.

The percentage of correct phonemes was calculated for each condition. Including a control condition with speech presented in isolation allowed us to isolate the specific impact of background noise on speech processing in PWS, if there is any.

The second task was a **dichotic aptitude** task. It allowed testing participants' binaural integration, binaural separation, and right ear advantage (Musiek & Weihing, 2011). Binaural integration and separation are crucial for perceiving complex auditory stimuli, such as speech (Broadbent, 1954). Dichotic aptitude tasks have also been used to demonstrate a typically superior ability to perceive auditory speech presented to the right ear. This phenomenon, known as the right ear advantage, is thought to reflect the left hemisphere's dominance for language processing (Kimura, 1961). Interestingly, atypical patterns of dichotic aptitude and right ear advantage have been observed in (some subgroups of) PWS (Foundas et al., 2004; Robb, Lynn, & O'Beirne, 2013). Here, dichotic aptitude was evaluated with 50 pairs of stimuli divided in five lists balanced acoustically, phonetically, linguistically, and semantically. Three of these lists were used in a free recall condition (in order: disyllabic nouns, two digits and monosyllabic adjectives), and two in a directed recall condition of the task (disyllabic nouns and three digits). In both cases, different stimuli were presented simultaneously to the participants' two ears. In the free recall condition, participants were asked to repeat what they heard. In the directed recall condition, they were asked to repeat only the stimulus presented in a designated ear. The designated ear changed every five pairs. All the stimuli were presented once to each ear. The percentage of words correctly repeated across the two conditions provided an index of dichotic listening ability . The percentage of words correctly repeated in the free and directed recall condition provided an index of dichotic integration and separation, respectively. Finally, an index of right ear advantage was measured by dividing the signed difference between the number of correct responses for the right and the left ear (right minus left) by the dichotic aptitude index.

The third task evaluated participants' ability to use interaural phase differences to detect acoustic signals that would be inaudible if presented to only one ear (Davidson et al., 2023; Hirsh, 1948). This phenomenon, known as the "masking level difference", reflects a **binaural advantage** in processing auditory signals in noise. We evaluated participants' masking level difference with 70 disyllabic nouns presented binaurally in noise (70 dB SPL). Stimuli were divided in seven lists. The lists were presented at decreasing levels of intensity,

from 70 to 40 dB SPL by steps of 5 dB SPL and in two conditions. In the homophasic condition, the stimuli and noise were presented in phase to both ears. In the antiphasic condition, the noise was presented in phase, but the stimuli were inter-aurally phase-shifted by 180 degrees. In both conditions, participants were asked to repeat all the words, presented one at a time, until they became unable to repeat correctly more than one word of a given list. The MLD was calculated as half the signed difference between the number of correctly repeated words in the two conditions (homophasic minus antiphasic).

The fourth task assessed participants' ability to perceive and correctly order multiple auditory signals – a skill known as "temporal ordering" (Chowsilpa, Bamiou, & Koohi, 2021; Hirsh, 1959; Koca et al., 2024; Mustek, Baran, & Pinheiro, 1990; Shinn, 2003). Previous research suggests that, on average, PWS may perform less efficiently than controls in temporal ordering tasks (Koca et al., 2024; Prestes et al., 2017), and that they may experience improvements in speech fluency following auditory temporal training programs that include temporal ordering exercises (Farazi et al., 2023). In the present study, temporal ordering was evaluated through two tasks: a **pitch pattern sequence** task and a **duration pattern sequence** task. In each trial, participants heard a series of three sounds presented binaurally, either varying in pitch (880 or 1122 Hz) or in duration (250 or 500 ms). Following the presentation, participants were asked to reproduce the pattern either verbally using the terms “high” and “low” (for the pitch pattern test) or “long” and “short” (for the duration pattern test), or by humming the sequence. Responses that inverted “high” and “low”, or “long” and “short” were accepted as correct provided that the overall pattern was correct, as in “low-high-low” in place of “high-low-high”. Each task consisted of 20 stimuli, and the percentage of correctly reproduced stimuli served as an index of the participant's temporal ordering abilities for both pitch and duration sequences.

2.3.5 Somatosensory abilities

Somatosensory information is widely recognized as critical for speech production (Ringel & Steer, 1963; Scott & Ringel, 1971; Tremblay, Shiller, & Ostry, 2003; see Guenther, 2016; Kent, 2024; Parrell & Houde, 2019 for recent reviews) and has been hypothesized to contribute to stuttering (Mysak, 1960; Van Riper, 1971). In this study, somatosensory testing encompassed assessments of participants’ orofacial proprioceptive and tactile abilities.

Proprioceptive testing included evaluations of effector position and movement amplitude, while tactile testing measured tactile point pressure sensitivity and force discrimination. Each of these assessments targeted different body parts, including speech-relevant orofacial structures (lips, tongue, jaw) and non-orofacial body parts (typically the hand, thumb, or wrist). Testing a non-orofacial body part served as a control condition, allowing us to examine the specificity, and to constrain hypotheses about the origin, of potential abnormalities in tongue and/or jaw somatosensation. AWS will be deemed to have an orofacial tactile or proprioceptive deficit only if: (1) they perform significantly below the controls' performance in a task assessing these abilities, and (2) they show a significantly larger difficulty to perform the task when it involves a speech-relevant orofacial structure as compared to their performance in a non-orofacial body part.

Participants' sense of hand, jaw and tongue position was assessed, in this order, in a **thickness discrimination** task. Stimuli were 14 Plexiglas "blocks" uniform in length (10 cm) and width (0.5 cm) but varying systematically in thickness, as illustrated in Figure 1A. Thickness ranged from 10 to 16 mm, with the following incremental differences: 1 mm steps from 10 to 12 mm, 0.5 mm steps from 12 to 15 mm, and finer 0.25 mm steps between 15 and 16 mm.

Each trial followed a standardized sequence: participants closed their eyes and were presented with two blocks—one "target" block (16 mm thick) and a thinner "foil" block—in randomized order. Afterward, they identified which block was thicker (first or second). The procedure was very similar in the three body parts tested.

In the HAND condition, the task was performed between the thumb and the index finger of the dominant hand. Participants placed their fist, little finger down, on a table in front of them, spreading their thumb and index finger as much as possible. The experimenter placed the first block on the palm of the participants' thumb. Participants were then instructed to pinch the block between their thumb and index finger and focus on its thickness for up to 5 seconds. After the block was removed, participants touched their thumb with their index finger, and the same procedure was repeated with the second block.

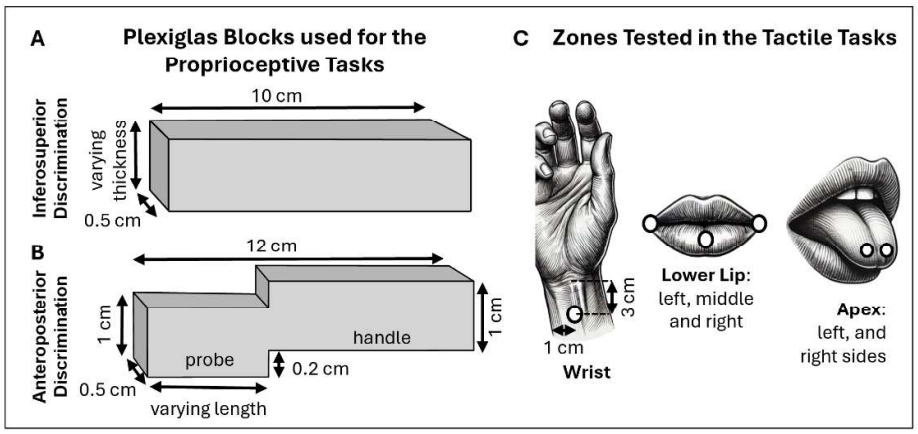
In the JAW condition, the task was performed between the upper and lower incisor. Participants had their mouths closed. The experimenter gently touched their lower lips with

the first block. Participants were asked to bite the block between their upper and lower incisors and focus on its thickness for up to 5 seconds. After the block was removed, participants closed their lips, and the procedure was repeated with the second block.

In the TONGUE condition, the task was performed between the upper incisors and the tongue. Participants' jaw was immobilized with a standardized dental silicone bite block (35 mm), placed between the right premolars. The experimenter positioned the upper edge of the first block against the lower ridge of the participants' upper incisors. Participants were asked to place the tip of their tongue on the lower surface of the block and focus on its width for up to 5 seconds. After the block was removed, participants lowered their tongue, and the same procedure was repeated for the second block.

Figure 1

Illustration of the material used in the proprioceptive tasks and of the zones tested in the tactile tasks



Participants performed the experiment three times for each of the three body parts tested. Each time, the stimuli were presented in order of increasing difficulty. In the first attempt, the starting difference between the two blocs was 3 mm for the hand and 2 mm for the jaw and tongue (foils with a thickness of 13 and 14 mm, respectively). If a participant failed the first trial, the starting difference was increased by 1 mm (foils with a thickness of 12 and 13 mm, respectively). Participants were required to make three consecutive correct responses to progress to the next level, in which the comparison block was the one most

similar to the target among the remaining options. This continued until the participant made an incorrect response, or until they made three consecutive correct responses at the smallest thickness difference of 0.25 mm, signaling the end of the attempt. The second and third attempts were similar, except that the starting difference corresponded to the level just before the level at which the previous attempt terminated. For instance, if attempt 1 terminated at a thickness difference of 0.5 mm, attempt 2 was initiated at a thickness difference of 0.75 mm. The thickness discrimination threshold was determined by averaging the thresholds estimated across the three attempts. Within each attempt, if the participant achieved three consecutive correct responses at the smallest thickness difference (0.25 mm), the threshold was recorded as 0.25 mm. Otherwise, the threshold corresponded to the level before the level at which an error occurred.

Participants' sense of thumb and tongue movement amplitude was assessed, in this order, in a **length discrimination** task. The stimuli were 14 Plexiglas "sticks," each comprising a longer section (the "handle") and a shorter section (the "probe"), offset by 5 mm, as shown in Figure 1B. The handle and probe had identical width (5 mm) and thickness (10 mm) but varied in length. Probe lengths ranged between 10 to 16 mm, increasing by 1 mm increments between 10 and 12 mm, by 0.5 mm increments between 12 and 15 mm, and by 0.25 mm increments between 15 and 16 mm.

Each trial followed a standardized sequence: participants closed their eyes and were presented with two sticks—the "target" stick who had a probe 16 mm long, and a "foil" one with a shorter probe—in randomized order. Afterward, they identified which stick had the longer probe (first or second). The procedure was very similar in the two body parts tested.

In the THUMB condition, participants placed the back of their opened hand on a table. The experimenter placed a first stick in the palm of the participant, who was invited to explore actively the length of the probe (i.e., between the end of the stick and the offset) by moving the thumb along the probe, going back and forth as many times as needed for up to 5 seconds. The experimenter removed the first stick, and the procedure was repeated for the second block.

In the TONGUE condition, participants were asked to stay still with their mouth closed. The experimenter gently touched the participant's lower lip with the stick. Participants

opened their mouths, and the experimenter then placed the upper offset of the first stick on their upper incisors. Participants were asked to explore actively the length of the probe (i.e., between the end of the stick and the offset) by moving the tongue along the probe, going back and forth as many times as needed for up to 5 seconds. The experimenter then removed the first block, and the procedure was repeated for the second block.

Again, participants performed the experiment three times for each body part. Each time, the stimuli were presented in order of increasing difficulty. In the first attempt, the starting difference between the two blocs was 3 and 2 mm for the thumb and tongue, respectively (foils with a probe length of 13 mm and 14 mm). If a participant failed the first trial, the starting difference was increased by 1 mm. Participants were required to make three consecutive correct responses to progress to the next level, where the comparison stick was the next more similar to the target one. This continued until the participant made an incorrect response, or until they made three consecutive correct responses at the smallest length difference of 0.25 mm, signaling the end of the attempt. The second and third attempts were similar, except that the starting difference corresponded to the level just before the level at which the previous attempt terminated. For instance, if attempt 1 terminated at a length difference of 0.5 mm, attempt 2 was initiated at a length difference of 0.75 mm. Length discrimination threshold was determined by averaging the thresholds estimated across the three attempts. Within each attempt, if the participant achieved three successive correct responses at the smallest length difference (0.25 mm), the threshold was recorded as 0.25 mm. Otherwise, the threshold corresponded to the level before that at which an error occurred.

As illustrated in Figure 1C, tactile point pressure sensitivity and force discrimination were examined on the wrists, on the midline of the lower lip (“middle lip”), on the right and left sides of the tongue apex, and on the right and left corners of the mouth (“left and right lip”), which are assumed to contribute critically to lip position estimation (Ito & Gomi, 2007; Shaiman & Gracco, 2002). These two tasks relied on a set of 17 von Frey filaments made of optical glass fibre (Fruhstorfer, Gross, & Selbmann, 2001) calibrated to produce contact forces ranging from 0.044 mN to 11.3 mN, with a logarithmic progression in steps of approximately $\sqrt{2}$. In both experiments, participants closed their eyes during the duration of the testing. The filaments were applied perpendicularly to the skin surface until bending

slightly for one second. In the WRIST condition, participants either selected to place their forearm on a table or on their thigh, palm down. For the LOWER LIP and the CORNERS OF THE MOUTH, participants staid still, mouth closed. For the APEX of the tongue, participants were asked to stick out the tongue at the beginning of each trial. For all the body parts except the tongue apex a small dot was placed on the stimulation area with a ballpoint pen to ensure a consistent placement of the filament across trials.

In each trial of the **point pressure sensitivity** task, participants were presented with a stimulation and were invited to signal the detection of a stimulation by saying "oui" ("yes" in French) for the wrists and by tapping on the table for the orofacial structures (to avoid damaging the filaments). Responses provided during or immediately after the stimulation (within approximately 1 second) were considered accurate detections. Successive stimulations were presented at a random interval ranging from 3 to 10 seconds.

Point pressure sensitivity was assessed three times at each location. Each time, the stimuli were presented in order of increasing difficulty. In the first attempt, the starting filament induced a pressure of 0.35 mN for the wrist and of 0.125 mN for the other body parts. If a participant failed the first trial, the next softer filament was used (filament with a force of 0.5 mN and 0.18 mN). Three successive detections of a given filament were followed by one (and up to three) trials with the next suppler filament, and so on. This continued until the participant made an incorrect response, or until they made three consecutive detections when stimulated with the softer filament (0.044 mN), signaling the end of the attempt. The second and third attempts were similar, except that the starting difference corresponded to the level just before the level at which the previous attempt terminated. For instance, if attempt 1 terminated with the filament applying a force of 0.044 mN, attempt 2 was initiated with a filament applying a force of 0.063 mN. Point pressure sensitivity threshold was determined by averaging the detection thresholds estimated across the three attempts. Within each attempt, if the participant achieved three consecutive correct detections of the softer filament (0.044 mN) the threshold was recorded as 0.044 mN. Otherwise, the threshold corresponded to the force of the filament before that at which an error occurred.

In each trial of the **force discrimination** task, the experimenter applied two different filaments in a random order: one filament producing a target force of 11.3 mN, and another filament producing a lighter force (the foil). The participant's task was to identify which of

the two intervals had the stronger contact force. Available foil forces were 8, 5.7, 4, 2.8, 2, 1.4 and 1 mN, which represent consecutive stimulation levels defined by the manufacturer on a logarithmic scale, consistent with the Weber–Fechner law.

Force discrimination was assessed three times at each location. Each time, the stimuli were presented in order of increasing difficulty. In the first attempt, the initial foil force was set at 4 mN. If a participant failed the first trial, the starting difference was increased (filament with a force of 2.8 mN). Three consecutive correct responses at a given foil force level were followed by one (and up to three) trials with a filament with a force level closer to that of the target. This continued until the participant made an incorrect response, or until they made three consecutive correct responses at the smallest possible force difference between the target and the foil (3.3 mN), signaling the end of the attempt. The second and third attempts were similar, except that the starting difference corresponded to the level just before the level at which the previous attempt terminated. For instance, if attempt 1 terminated at a force difference of 3.3 mN, attempt 2 was initiated at a force difference of 5.6 mN. The second and third attempts followed the same procedure, except that the starting foil force was set to the level just less similar than the one at which the previous attempt had terminated. To estimate the force discrimination thresholds, we first determined the threshold estimated at each attempt. Within each attempt, if the participant achieved three consecutive correct responses at the smallest force difference, the threshold was recorded as 3.3 mN. Otherwise, the threshold corresponded to the force difference at the level just before that at which an error occurred. Then, we converted this threshold in decibel units relative to the target force ($10\log_{10}(F/F0)$ where F is the force level of interest, and F0 is the target force of 11.3 mN. Finally, the force discrimination thresholds was calculated as the mean of the decibel values of the estimated thresholds of the three attempts.

2.3.6 Rhythmic/timing abilities

Participant's rhythmic and timing abilities were evaluated using a **synchronization** and **continuation** task. In this task, participants are first asked to synchronize their movements (e.g., tapping or speaking) with an external rhythmic cue, such as a metronome. After this synchronization phase, the rhythmic cues stop and participants are asked to continue producing the rhythm on their own, maintaining the timing of the original cues – the continuation phase. This task is designed to assess timing and rhythmic abilities that are crucial for fluent speech production. The

synchronization phase evaluates how well participants can coordinate their motor output (producing the movement) with the auditory input (hearing the cue). And the continuation phase assesses the ability to sustain a consistent rhythm under the control of internal timing mechanisms (without external guidance). Previous studies have reported that, on average, groups of PWS tend to be more variable than controls during the synchronization phase, and/or less accurate in the continuation task, pointing to possible sensorimotor and timing deficits (Max & Yudman, 2003; Olander et al., 2010).

A verbal and a manual modality of the task was presented to the participants. The verbal modality was designed to test speech-related orofacial movements (pronouncing the syllable /ta/) and the manual modality was designed to test finger movements (a key press with the index finger). Again, the inclusion of the manual modality served as a control condition to determine whether any observed rhythmic difficulties in the verbal modality were specific to speech-related tasks or whether they might reflect more general rhythmic deficits or other task-related factors.

Each trial began with a preparation cue—a visual progress indicator (a 'loading bar') filling up to represent the passage of time—displayed on the screen for 3 seconds. Following this, participants heard a periodic sequence of 28 auditory metronome-like ticks presented at 4 Hz (with each tick lasting 42 ms) and at a comfortable volume through loudspeakers. Participants were instructed to synchronize their movements as accurately as possible with the auditory ticks during the synchronization phase. After the auditory cues ceased, participants were asked to continue their movements at the same tempo for an additional 5 seconds during the continuation phase. Participants were encouraged to begin moving in synchrony with the tones as soon as possible but were informed that their performance would be evaluated only following the word "GO"— displayed on the monitor two seconds after the onset of the first tone (immediately after the 8^{th} tone).

Each participant completed a total of nine trials for each of the two modalities, with the first trial in each modality serving as practice trial. The two modalities were always presented in the same order, starting with the verbal modality. In the manual modality, the timing of key presses was recorded using PsychoPy software. In the verbal modality, speech was audio-recorded, and the timing of the syllable nuclei was extracted using a combination of automated and manual analyses. First, a PRAAT script (Jong & Wempe, 2009) was used to identify the syllable nuclei. The script employed (1) intensity analysis to detect energy peaks, and (2) voicedness analysis to exclude

unvoiced peaks. All automated detections were subsequently manually verified for accuracy by the first author.

Performance was quantified with two measures: accuracy and precision. Accuracy was calculated as the mean, across all eight trials of the absolute difference between the expected frequency (4 Hz) and participants' actual movement frequency. Precision, indicating the consistency of movement timing, was calculated as the mean, across all eight trials of the standard deviation of the intervals between consecutive movements within each trial.

2.3.7 Motor abilities

According to the speech motor skills hypothesis, PWS exhibit limitations in motor control, characterized notably by increased variability and slower execution of movements (Namasivayam & Lieshout, 2011). Here, participants' motor skills were evaluated through a task designed to measure participant's **maximum speed** of "simple" and "alternating" verbal, manual and orofacial movements. The verbal modality required participants to repeat the syllable (/po/) in the 'simple' condition and the multisyllabic sequence (/potu/) in the 'alternating' condition. The manual modality involved single key presses (down arrow key) with the right index in the 'simple' condition and alternating between two key presses (up and down arrow keys) with the right index in the 'alternating' condition. For the orofacial modality, participants performed repetitive lip-smacking movements in the 'simple' condition, while the 'alternating' condition involved alternating between lip-smacking and tongue-clicking. As in the other experiments describe above, the inclusion of the manual and the orofacial modalities served as control conditions, enabling to differentiate between putative abnormalities specific to speech-related orofacial movements those affecting orofacial movements more generally, and those resulting from more general motor or rhythmic deficits, attentional limitations, or other task-related factors.

Each trial of the task began with a visual countdown displayed on a computer screen for 5 seconds, followed by an auditory tone (440 Hz, 500 ms) delivered through loudspeakers to signal the start of the trial. Participants were instructed to perform the specified motor task as quickly as possible during the subsequent 5-second interval. The trial concluded with a second auditory tone (440 Hz, 500 ms) marking the end of the task.

Participants completed one practice trial followed by five experimental trials for each modality. The modalities were presented in a fixed order: first, the verbal modality; next, the orofacial modality; and finally, the manual modality, which was performed with the left hand and then with the right hand. For each modality, participants began with the simple condition and progressed to the alternating one.

For the manual modality, the timing of key presses was recorded using PsychoPy software. For the verbal and orofacial modalities, audio recordings of the participants' productions were captured. The timing of each movement was extracted from these recordings using a PRAAT script (Jong & Wempe, 2009), which identifies events based on peaks in intensity and voicedness. All automatic extractions were subsequently reviewed and verified for accuracy by the first author.

Movement performance in each condition (simple and alternating) and modality was evaluated using two key metrics of execution proficiency: maximal execution rate and variability of inter-movement intervals. Maximal execution rate was defined as the total number of movements executed across all five trials, divided by the cumulative duration of the movement phases (5 seconds per trial). This measure provided an average execution rate in movements per second. Variability of inter-movement intervals was assessed by calculating the average standard deviation of the intervals between consecutive movements within each trial, aggregated across all five trials. This metric served as an indicator of temporal consistency during task performance. These measures enabled a comprehensive assessment of both the speed and regularity of motor execution across the verbal, orofacial and manual modalities.

3. Analyses and results

The organized raw data and the code used to run the analyses and generate the figures have been made available on OSF (https://osf.io/z84nx/).

3.1 General analyses

3.1.1 Logic guiding the analysis and interpretation of the data

The objective of this study was to investigate the patterns of strengths and weaknesses in 12 AWS across five major ability domains in order to identify and evaluate potential causal factors underlying each participant's stuttering.

Data interpretation will be guided by a conservative approach in ruling out potential causal factors (i.e. that minimizes the likelihood of Type II errors). Both clinically and theoretically, we consider it particularly important to avoid mistakenly concluding that a factor does not contribute to stuttering when it actually does, as such errors could prematurely eliminate a factor from further investigation. Hence, below, a factor (e.g., a semantic processing disorder) will be ruled out as a potential contributor to an AWS' stuttering only when that AWS performs above the controls' median performance in a task assessing that factor. In contrast, performances within the normal range but below the median of the controls will receive less emphasis, as they offer a less definitive basis for interpretation or conclusions.

In addition, our approach to identifying potential causal factors for an AWS's stuttering will be guided by the recognition that the origin of abnormally low performance in any given task, modality or condition cannot be determined in isolation. Every task engages multiple abilities, making the origin of poor performance inherently multidetermined. For example, poor performance on a task designed to assess semantic processing might indicate a semantic processing disorder, but it could also arise from other challenges, such as impairments in object recognition (if the task involves visual stimuli), decision-making, response planning, or execution. More general factors like anxiety or motivation may also influence performance. For that reason, whenever possible, we will rely on two key criteria to conclude that a particular ability might contribute to an AWS's stuttering: (1) the AWS's performance must be significantly below that of the control participants in a task assessing the ability, and (2) their performance must also be significantly worse than their own performance in a related control task or condition. For example, abnormally low performance on a semantic task combined with significantly better performance on an object decision task would suggest that a semantic weakness might underlie the AWS's stuttering. In contrast, significantly impaired performance on both a speech-related task and its control counterpart will be given less

emphasis, as such results offer a less definitive basis for interpretation. A consequence of this rationale is that although the results of control tasks and conditions will be systematically presented in the result section, these will only be discussed when they provide insight into the findings related to speech tasks and abilities.

3.1.2 Criteria for the definition of weaknesses, strengths and dissociation

The general logic described above required to establish clear criteria for weaknesses, strengths and dissociation of performance. We decided to classify each AWS' performance in one of three categories based on the following criteria: (1) performances below the fifth percentile were considered "weaknesses"; (2) performances between the fifth percentile and the median were deemed "normal"; and (3) performances numerically above the control group's median were considered "strengths."

We estimated percentile thresholds based on the following two approaches. For most tasks we used Crawford and Howell's modified *t* test (1998), from which the *p* value obtained for a score provides an estimate of the proportion of controls with performance below that score (Crawford & Garthwaite, 2006). Compared to the more traditional z-score approach.

Crawford and Howell's modified *t* test has the advantage of controlling the Type I error rate regardless of the size of the control sample, even in the face of severe skew and/or leptokurtosis (Crawford et al., 2006). In some tasks however, in which the dispersion of control scores was near-zero (due to ceiling effects), Crawford and Howell's approach was inappropriate, and we opted for an estimate of percentile thresholds based on linear interpolation from the empirical cumulative control distribution (Hyndman & Fan 1996), as implemented by the default method in the "quantile" function of R's stats package (R Core Team, 2022). Concretely, we adopted this second approach whenever the interquartile range (IQR) of the control distribution was equal to zero. This occurred for two tasks within the auditory ability domain (pitch pattern sequence task and speech-in-noise recognition task) and for four body parts in the point pressure sensitivity task within the tactile ability domain (left and right corners of the mouth and left and right sides of the tongue apex).

In turn, we used Crawford's Revised Standardized Difference Test (RSDT; Crawford & Garthwaite, 2005) to investigate dissociations between performance on speech-related

versus control tasks and conditions. The RSDT tests whether a difference in performance between two conditions in an AWS exceeds what would be expected based on the control group's performance in the same two tasks or conditions (i.e., a test for an interaction when one of the group is of size N=1). As for *t* tests, we did not compute RSDT when the interquartile range of the control scores was zero in at least one of the two conditions to be compared.

3.2 Group level analyses

Although this study was conceived as a multiple single-case design, we will also provide traditional group-level comparisons for each measure, tasks and conditions. However, we urge caution in interpreting these group-level comparisons, as they may be affected by the limitations of low statistical power due to our small sample size. As a matter of caution, we hence restricted group-level analyses to paired *t* test comparisons and did not explore interactions between group and tasks, modalities or conditions.

Group-level comparisons were performed with Welch's unequal variance *t* test to account for differences in variances and sample sizes between groups (R Core Team, 2022). Although *p* values from group-level *t* tests cannot be considered indicators of effect size (unlike *p* values from Crawford and Howell's modified *t* test), for the sake of simplicity in figures and notations, group differences were classified as a "weakness" when the group of PWS performed significantly below the control group at $p < .05$.

3.3 Outlier detection and exclusion

In some tasks, one or more control participants exhibited performance levels that appeared abnormally high or low. While the exclusion of "outliers" is a common practice in the field, this process is somewhat arbitrary. Here, we opted for a systematic exclusion of outliers in all tasks based on the following rules. For most of the tasks, "outliers" were identified and excluded based on the 1.5 x IQR rule, i.e., performances falling outside a range between 1.5 times the interquartile range (IQR) below the first quartile and 1.5 times the interquartile range above the 3rd quartile. For tasks where the IQR was 0 however, we systematically removed the worst performing participant. When the worst performance level was achieved by more than one participant no data was removed. In the interest of

transparency, all data points excluded based on these criteria are marked with filled triangles or diamonds in all the figures presenting the results. In the interest of robustness, we also conducted all analyses without excluding any outliers (see supplemental materials). The main results and conclusions remained unchanged (see general discussion).

3.4 Missing data

Data from two AWS were missing in the verbal modality of the inhibitory control assessment due to a technical issue. The "phonetic task" of the linguistic assessment was interrupted for one control participant who had clearly misunderstood the instructions. Another control participant opted out of performing the orofacial modality of the motor abilities assessment. As per our ethical agreement, the participant was not asked to justify this decision. Another control participant opted out of the proprioceptive assessments of the tongue and jaw, evoking a painful temporo-mandibular joint at the time of testing. We decided not to assess that participant's hand proprioception.

4. Results

4.1 Linguistic assessments: efficiency of access to successive levels of representation

Figure 2 presents participants' efficiency and transitioning costs across the six tasks of the linguistic assessments (see also Figure S1 for accuracy and response latencies, and Figure S2 for the results of the experiments analyzed without excluding outliers). At the group level, AWS exhibited a weakness in object recognition ($t(14.5) = 2.2$, $p = .02$), a weakness in accessing semantic information ($t(16.92) = 2.49$, $p = .012$), and a weakness in transitioning between these two levels ($t(14.73) = 3.1$, $p = .004$).

Despite these group-level findings, individual profiles were markedly heterogeneous. Consistent with the group results, AWS 1, 2, 6, and 7 exhibited significant weaknesses in object recognition, in semantic processing and in transitioning between these two levels. The origin of the object recognition weakness (both for these individuals and the group as a whole) remains indeterminate, as it could reflect a true recognition impairment or the influence of broader task-related factors (e.g., attention, response planning, or execution).

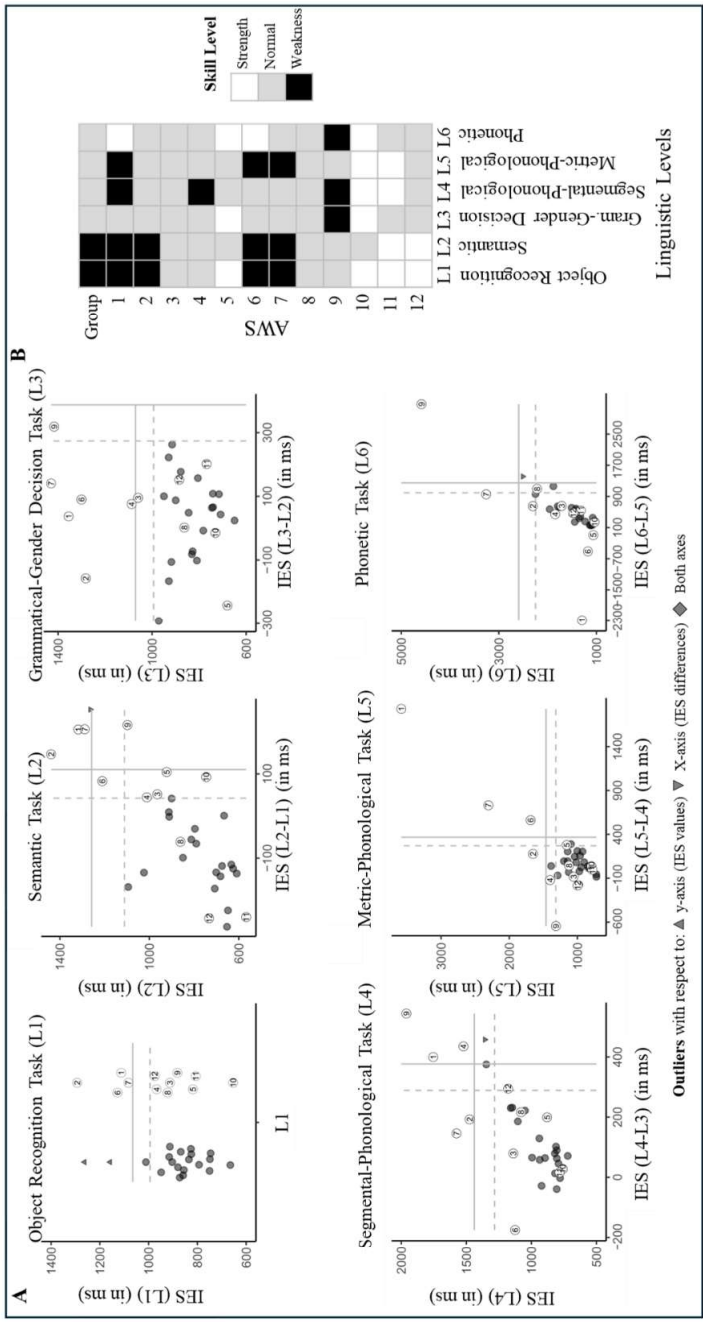
However, the combination of semantic processing weaknesses and abnormally high semantic transitioning costs in these individuals provides compelling evidence for an underlying semantic disorder. This finding aligns with prior research suggesting a relationship between stuttering and reduced semantic efficiency (Bosshardt & Fransen, 1996). However, no such weaknesses were observed in the remaining eight AWS. Notably, AWS 11 and 12 performed better than the average control participants in both semantic efficiency and transitioning costs (Figure 2B).

Single-case analyses revealed additional heterogeneity across the other tasks. At least one AWS struggled in both efficiency and transitioning cost in all the other language tasks/levels. AWS 9 exhibited a weakness in retrieving grammatical gender information. AWS 1, 4, and 9 struggled with segmental-phonological processing. AWS 1, 6, and 7 struggled with metric-phonological information, and AWS 9 had a weakness in phonetic/articulatory processing. However, other AWS performed well above the control group average in both efficiency and transitioning costs across these same tasks. This suggests that these AWS' stuttering is unlikely to stem from a degree of efficiency below that found in the general population in the type of linguistic skills and abilities assessed in these tasks. Overall, half of the AWS participants showed no evidence of weakness in any task of the language tasks administered.

These contrasting findings underscore the necessity of case-by-case analysis. While the results suggest that stuttering is associated with various types of psycholinguistic weaknesses, these weaknesses appear to differ across individuals.

Figure 2

Results of the linguistic assessment



Note. A. Scatter plots of performances for the six tasks of the linguistic assessments are shown for controls (grey filled circles, triangles indicating outliers) and AWS (white circles labeled with their ID). Grey lines (vertical and horizontal) indicate statistical thresholds, corresponding to the first (plain; p = 0.01) and fifth (dashed; p = 0.05) estimated percentiles of the control distribution for each axis. B. Overall performance of each participant in all the tasks administered obtained from the combined performances in efficiency and transitioning cost. As detailed in the text, performances of AWS exceeding the fifth percentile threshold along both axes (i.e. falling within the upper right quadrant defined by the dashed lines) are considered as a weakness, performances of AWS numerically below the controls' median along both axes indicate a strength. In all other cases, the performance is classified as normal.

4.2 Linguistic assessment: inhibitory control

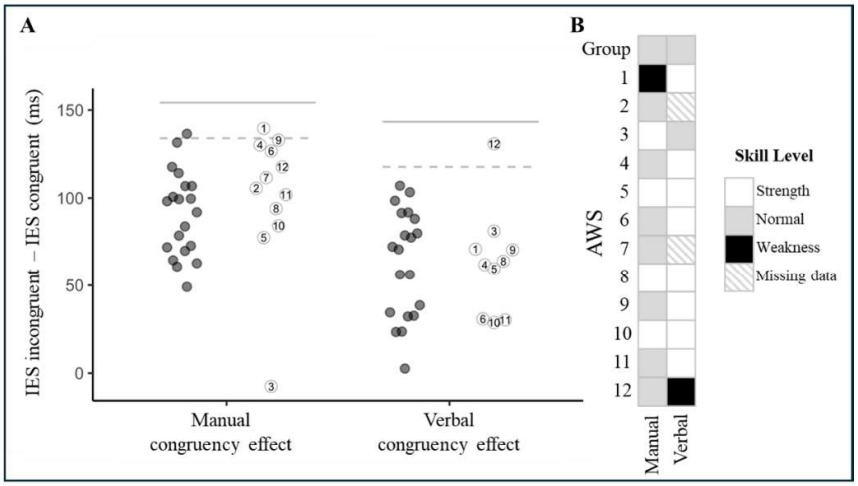
The main results of the inhibitory control assessment are presented in Figure 3 (see also Figure S3 for the accuracy, response latencies, and inverse efficiency scores for each modality). There was no significant group difference in either the verbal ($t(18.34) = -.01$, $p = .5$) or the manual ($t(16.11) = .82$, $p = .21$) congruency effect.

Single-case analyses largely corroborated the group results. Most of the AWS showed no sign of a larger congruency effect than the controls in any of the modality. In fact, eight of the 10 AWS whose data could be analyzed (see missing data section) showed a strength in the verbal modality. This finding suggests that these AWS' stuttering is unlikely to stem from a degree of efficiency below that found in the general population in the type of speech-related inhibitory control ability assessed in this task.

One of the AWS (AWS 12) showed a significantly larger congruency effect than the controls in the verbal modality, suggesting that speech-related inhibitory control difficulty is associated with stuttering in at least a small proportion of the AWS. However, this finding must be interpreted with caution as this deficit was not significantly larger in the verbal than in the manual modality (RSDT, t (19) = 1.09, p = .14). This lack of specificity makes it difficult to rule out that the observed deficit may arise from broader task-related factors rather than from speech-related inhibitory control difficulty specifically.

Figure 3

Results of inhibitory control assessment



Note. A. Congruency effect for controls (grey filled circles) and AWS (white circles labeled with their ID) for the manual and verbal modalities. Grey lines indicate statistical thresholds, corresponding to the first (plain) and fifth (dashed) estimated percentiles of the control distribution. B. Summary of the results. As detailed in the text, performances of AWS exceeding the fifth percentile threshold are considered as a weakness, performances of AWS numerically below the controls' median indicate a strength. In all other cases, the performance is classified as normal.

4.3 Auditory abilities

The main results of the auditory assessments are presented in Figure 4 (see also Figure S4 for the results of the pitch and duration pattern sequence tasks analyzed without excluding outliers). At the group level, the AWS demonstrated a weakness in dichotic listening ($t(13.99) = -2.3$, $p = .019$), driven by a weakness in dichotic integration ($t(13.54) = -1.8$, $p = .047$). No other group-level differences were observed (all other $ts < 1.7$, $ps > .055$).

Consistent with the group findings, single-case analyses revealed that AWS 2, 4, 6, 9, and 12 exhibited weaknesses in dichotic listening and integration. This result aligns with previous research (Foundas et al., 2004; Robb et al., 2013). However, these weaknesses were not universal among AWS. The seven other AWS showed no sign of dichotic aptitude

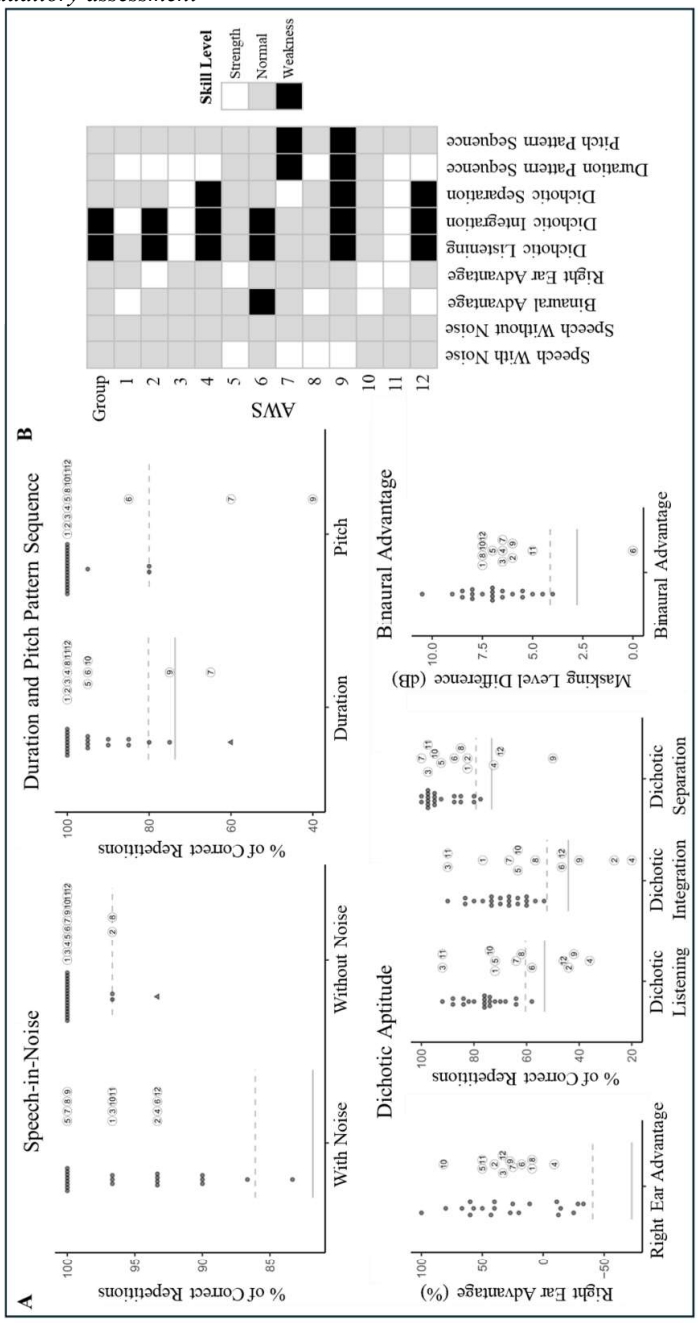
weaknesses and for four of them dichotic integration and/or separation was a strength (AWS 1, 3, 7, and 11).

Further insights emerged from single-case analyses. Some AWS performed above the median level of controls on all tasks where the median performance was not at ceiling. This suggests that for at least these individuals, stuttering is unlikely to stem from a degree of efficiency below that found in the general population in the skills and abilities assessed in these evaluations. Conversely, in other cases, the analyses highlighted additional potential auditory contributors to stuttering. AWS 4, 9, and 12 exhibited a deficit in dichotic separation. AWS 7 and 9 exhibited difficulties in both duration and pitch pattern sequence *d*iscrimination, and AWS 6 showed no sign of the typical binaural advantage in masking level difference.

Overall, these results suggest that different types of auditory weaknesses are associated with stuttering in different AWS, but they also highlight that auditory disorders in AWS are neither uniform nor universal.

Figure 4

Results of auditory assessment



Note. A. Performance of controls (grey filled circles, triangles indicating outliers) and AWS (white circles labeled with their ID) in the auditory assessment. Grey lines indicate statistical thresholds, corresponding to the first (plain) and fifth (dashed) estimated percentiles of the control distribution. B. Summary of the results of the AWS. Performances below the fifth percentile threshold are considered as a weakness, performances of AWS numerically above the controls' median indicate a strength. In all other cases, the performance is classified as normal.

4.4 Somatosensory skills

The main results of the proprioceptive and tactile assessments are presented in Figure 5 and 6, respectively (see also Figure S5 and S6 for the results of these assessments analyzed without excluding outliers). At the group level, the AWS demonstrated significant difficulties in the two tasks for most body parts. In the proprioceptive assessment, they exhibited a deficit in hand thickness discrimination ($t(26.36) = 4.03$, $p < .001$), as well as a weakness in tongue thickness discrimination ($t(23.06) = 2.15$, $p = .02$) and tongue length discrimination ($t(26.27) = 2.24$, $p = .017$). In the tactile assessments, they exhibited difficulties in detecting point pressure across all tested body parts (all ts > 2.29, all $ps < .05$), except for the right wrist ($t(11.18) = 1.68$, $p < .06$) , as well as in discriminating between forces in most tested body parts (all ts > 2.29, all $ps < .05$), with the exception of the right wrist ($t(17.11) = -1.67$, $p < .057$) and left corner of the mouth ($t(13.5) = -.98$, $p < .17$).

Despite these group-level findings, individual proprioceptive profiles were heterogeneous, and the comparisons between the AWS' performance for the speech-related orofacial structures and control body parts provided only limited support for the hypothesis that stuttering may be linked to reduced orofacial proprioceptive ability. Consistent with group results, six AWS (AWS 1, 2, 5, 6, 7, and 9) showed weaknesses in tongue and/or jaw thickness discrimination task, with AWS 6 also displaying weaknesses in the tongue length discrimination task. In most of these cases, however, these AWS (AWS 1, 2, 5, 6, and 9) did not exhibit a statistically significant difference between their performance in speech-related orofacial structures (tongue or jaw) and control body parts (all RSDT ps > .3), making it difficult to know whether these weaknesses are due to a proprioceptive disorder per se, or to other task-related factors.

Furthermore, it is noteworthy that six other AWS displayed no weaknesses in any proprioceptive task, and that several performed above the control group average in some (AWS 3, 5, 8, 12) or even in all tasks (AWS 11), suggesting that these AWS' stuttering is unlikely to stem from a degree of efficiency below that found in the general population in the type of orofacial proprioceptive skills assessed in these tasks. The only exception, and clear evidence of an orofacial proprioceptive disorder concerned AWS 7, whose performance in the jaw thickness discrimination task was also significantly worse than his performance in the related control condition (RSDT, $t(17) = 2.56$, $p = 0.019$). Hence, these findings collectively provide clear evidence for a link between stuttering and reduced orofacial proprioceptive ability in only one AWS (AWS 7).

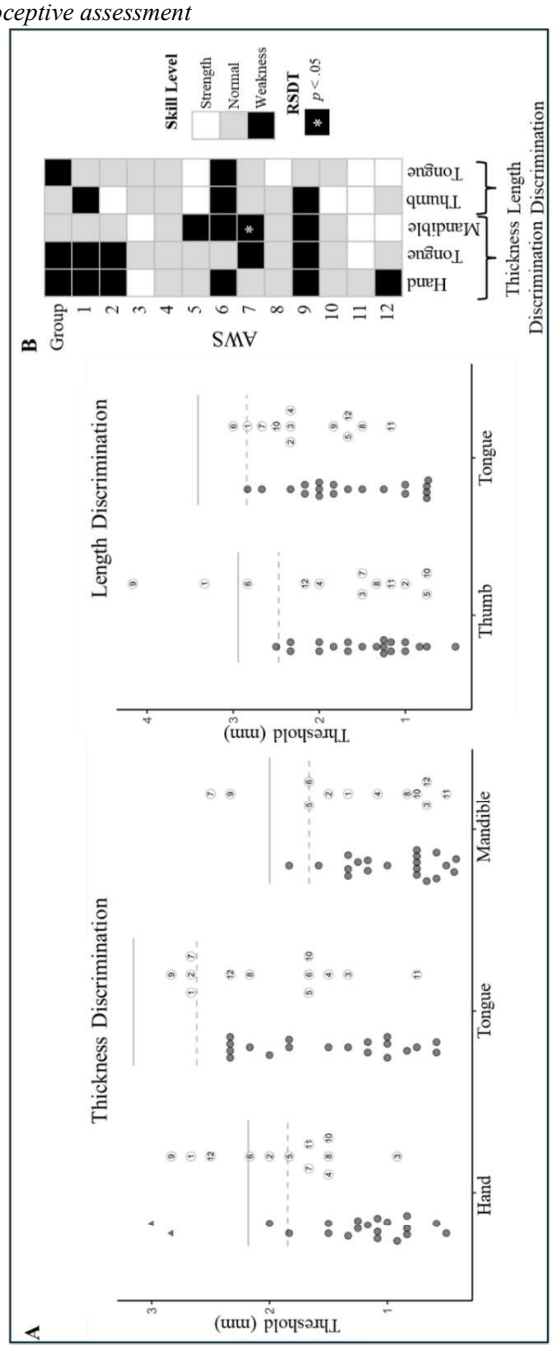
The results of the tactile assessment showed markedly more consistent results. All the AWS exhibited some weaknesses in at least one of the two tasks. AWS 11 had an abnormally high threshold for detecting point pressure at both corners of the mouth. AWS 5 displayed an elevated threshold for force discrimination on the lower lip and right apex of the tongue. The remaining 10 AWS all exhibited abnormally high thresholds in both point pressure sensitivity and force discrimination. In addition, and importantly, in most of these cases for the force discrimination tasks, these weaknesses are unlikely to result from general task-related demands, as the AWS performances in orofacial body parts were often significantly worse than in control areas (see Figure 6B). Indeed, all the AWS demonstrated at least one clear weakness in an orofacial area that was also associated with a significantly worse performance than in the corresponding control body part. This strongly suggests that the observed weaknesses reflect specific tactile impairments rather than nonspecific influences such as attention or task comprehension. RSDT comparisons were not conducted for the point pressure sensitivity task, as the control group's interquartile ranges for the Lip and Apex regions were zero. Nonetheless, it is worth noting that 9 of the 11/12 AWS who presented with a weakness in at least one orofacial area had a normal performance in a control condition.

It may be worth concluding this section by briefly addressing the two participants in our sample who presented with atypical orofacial conditions. To our knowledge, there is no physiological or empirical basis to expect that AWS 8's history of cleft palate would affect tactile sensitivity of the lower lip or tongue. Consistent with this, his performance in the point

pressure sensitivity task showed relatively better results for the lips and tongue compared to the control region (the wrists). Similarly, the fact that AWS 2 and AWS 8 both performed within the typical range in the *Jaw* condition of the thickness discrimination task, which requires discrimination of small interdental distances, suggests that their dental prostheses did not interfere with proprioceptive processing in this context. These observations support the interpretation that the findings reported above are not artifacts of including participants with orofacial anatomical anomalies.

Figure 5

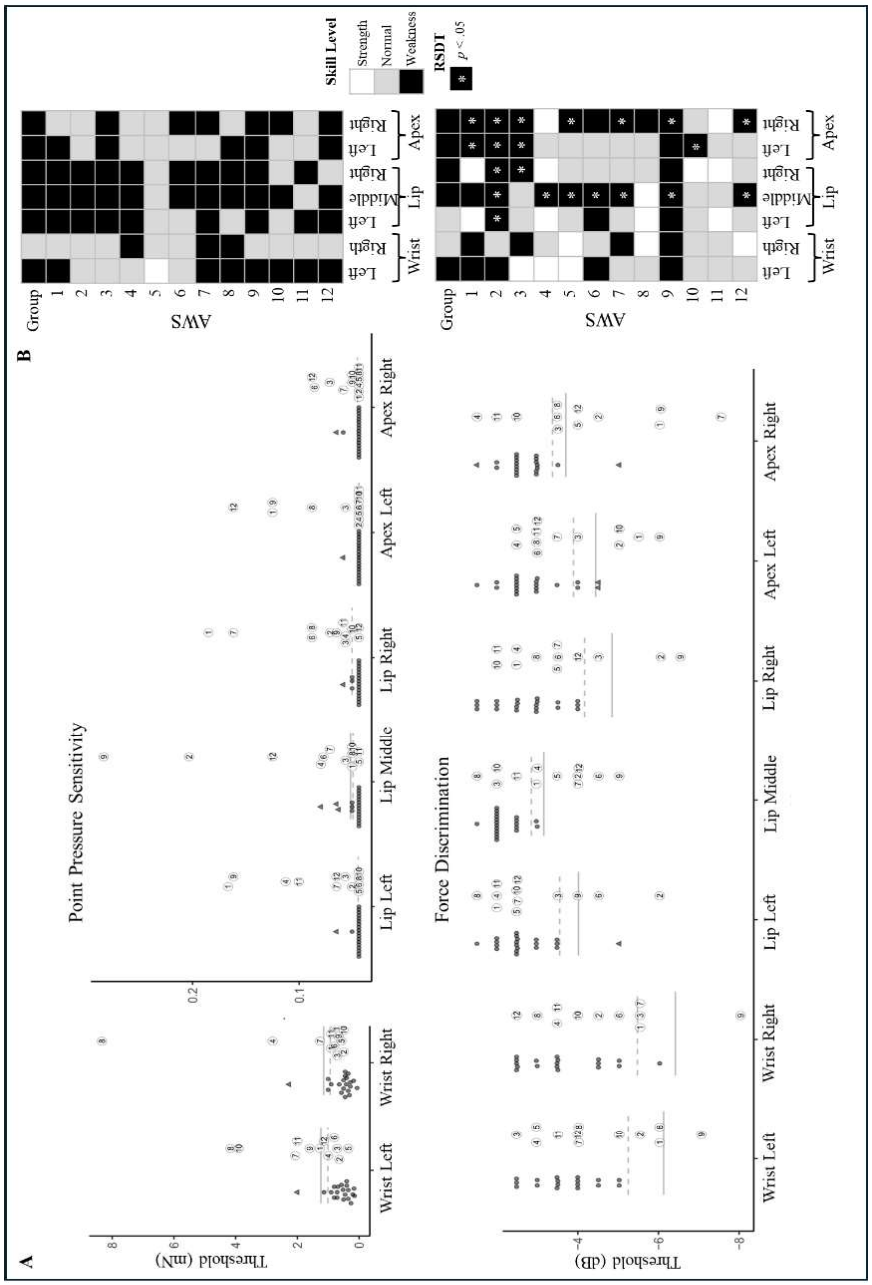
Results of proprioceptive assessment



Note. A. Performance of controls (grey filled circles, triangles indicating outliers) and AWS (white circles labeled with their ID) in proprioceptive assessments. Grey lines indicate statistical thresholds, corresponding to the first (plain) and fifth (dashed) estimated percentiles of the control distribution. B. Summary of the results. Performances below the fifth percentile threshold are considered as a weakness, performances of AWS numerically above the controls' median indicate a strength. In all other cases, the performance is classified as normal. Cases displaying weaknesses in speech related orofacial structures (black cells for Tongue or the Mandible) were compared to their respective performances in control body parts (Hand and Thumb). The * symbol indicates when a significant difference ($p < .05$) was obtained in at least one of the corresponding RSDT tests (Crawford & Garthwaite, 2005).

Figure 6

Results of tactile assessment



Note. A. Performance of controls (grey filled circles, triangles indicating outliers) and AWS (white circles labeled with their ID) in the tactile assessments. Grey lines indicate statistical thresholds, corresponding to the first (plain) and fifth (dashed) estimated percentiles of the control distribution. B. Summary of the results. Performances below the fifth percentile threshold are considered as a weakness, performances of AWS numerically above the controls' median indicate a strength. In all other cases, the performance is classified as normal. Cases displaying a weakness in speech related orofacial structures (black cells in Lip and Apex regions) were compared to their respective performances in control body parts (Left and Right Wrists). The * symbol indicates when a significant difference ($p < .05$) was obtained in at least one of the two corresponding RSDT tests (Crawford & Garthwaite, 2005). Note that for Point Pressure Sensitivity, RSDT comparisons were not conducted, as the control group's interquartile ranges for the Lip and Apex regions were zero.

4.5 Rhythmic/timing skills

The main results of the rhythmic/timing assessment are presented in Figure 7 (see also Figure S7 for the results of these tests analyzed without excluding outliers). At the group level, the AWS exhibited a significant weakness in verbal synchronization precision ($t(11.87) = 1.98$, $p = .036$). No other group difference was statistically significant (all $ts < 1.54$, all $ps > .073$).

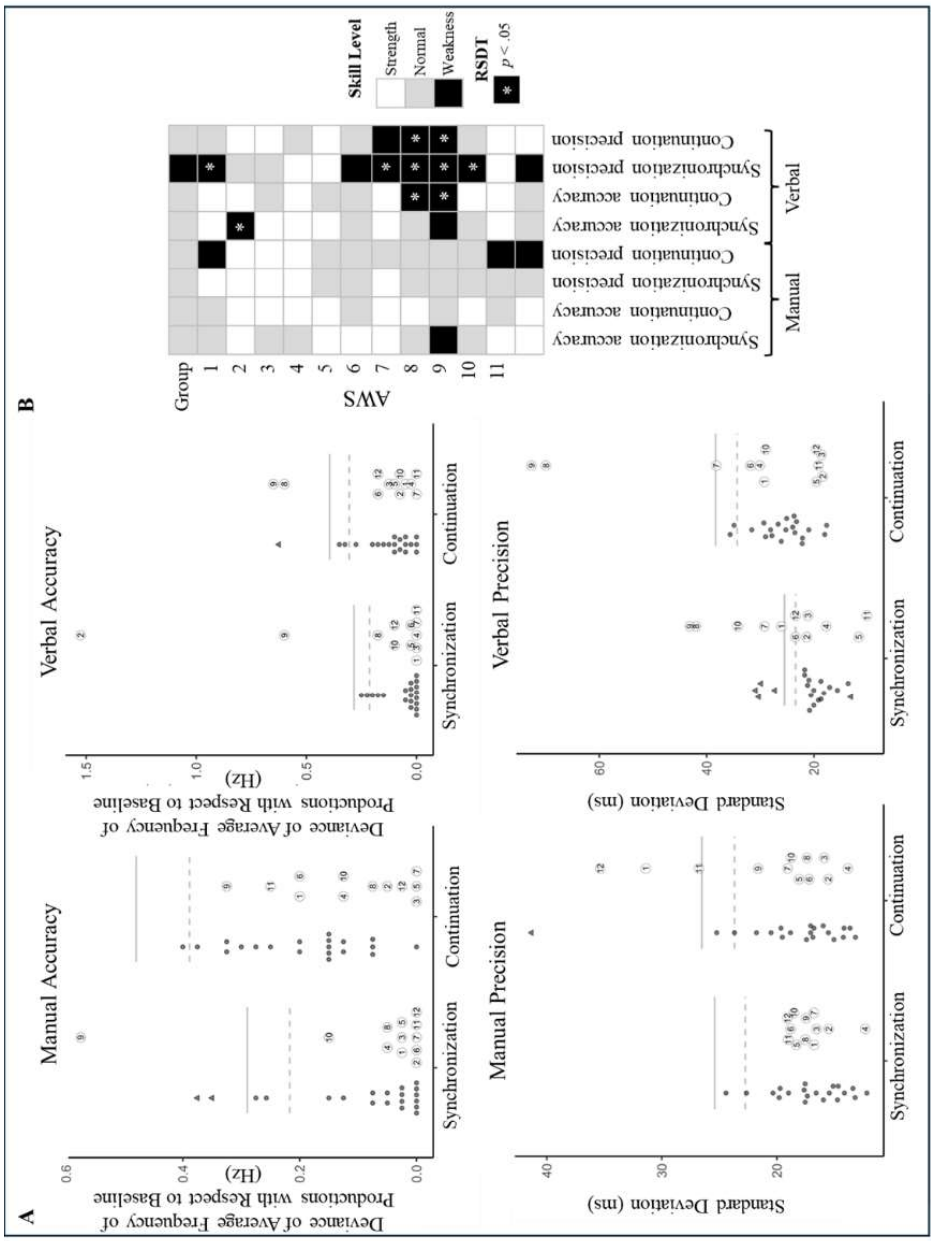
Consistent with the group findings, single-case analyses revealed that seven AWS (AWS 1, 6-10 and 12) exhibited significant weaknesses in verbal synchronization precision, that is, significantly more variable production of the syllable specifically during the synchronization phase. This trend was not observed in the manual modality, where all participants were well within the range of the controls in the synchronization phase. In three of these AWS (AWS 1, 7 and 10), the performance in the verbal synchronization precision was also significantly worse than in the manual synchronization precision (both RSDT $ps < .05$, all other RSDT $ps > .24$). This provided clear evidence of a synchronization disorder in at least these three AWS.

This said, it is also interesting to note that other AWS performed at a normal level of performance (AWS 2 and 3), or even at an above average (strength) level of performance in verbal synchronization precision (AWS 4, 5 and 11). This suggests that these AWS' stuttering is unlikely

to stem from a degree of efficiency below that found in the general population in the type of rhythmic and timing skills assessed in these tasks. This places clear limits on the generalization of this finding within the population of AWS suggesting that a synchronization precision disorder may affect some, but not all AWS. It seems equally important not to limit the association of stuttering with verbal precision during the synchronization phase, as several other AWS also showed clear signs of other types of weaknesses affecting all the possible index of performance in the verbal modality.

Figure 7

Results of rhythmic assessment



Note. A. Performances of controls (grey filled circles, triangles indicating outliers) and AWS (white circles labeled with their ID) are shown for the two measures of rhythmic abilities, accuracy of the frequency production from the expected 4 Hz baseline (upper panels) and precision of production (lower panels), for finger movement (left panels) and speech-related movement (right panels) tasks. Grey lines indicate statistical thresholds, corresponding to the first (plain) and fifth (dashed) estimated percentiles of the control distribution. B. Summary of the results. Performances below the fifth percentile threshold are considered as a weakness, performances of AWS numerically above the controls' median indicate a strength. In all other cases, the performance is classified as normal. Cases displaying difficulties in speech related zones (dark cells in Verbal modality) were compared to their respective performances in the associated control condition (Manual modality). The * symbol indicates when a significant difference ($p < .05$) was obtained in the corresponding RSDT test (Crawford & Garthwaite, 2005).

4.6 Motor abilities

The main results are of the motor assessments presented in Figure 8 (see also Figure S8). As a group, the AWS were significantly slower than the control group for the simple ($t(29.91) = -2.34$, $p = .013$) and alternating ($t(29.53) = -2.73$, $p = .005$) conditions of the verbal modality.

At first sight, it may be tempting to conclude that stuttering may emerge from an insufficiently efficient speech-related motor system. The individual profiles were at odds with the group results, however. No AWS was significantly slower than controls for either the simple or the alternating condition of the verbal modality. Some of them even performed numerically above the control group's median ("strength") in these two conditions (AWS 5, 10 and 11). This suggests that these AWS' stuttering is unlikely to stem from a degree of efficiency below that found in the general population in the type of motor abilities assessed in these tasks.

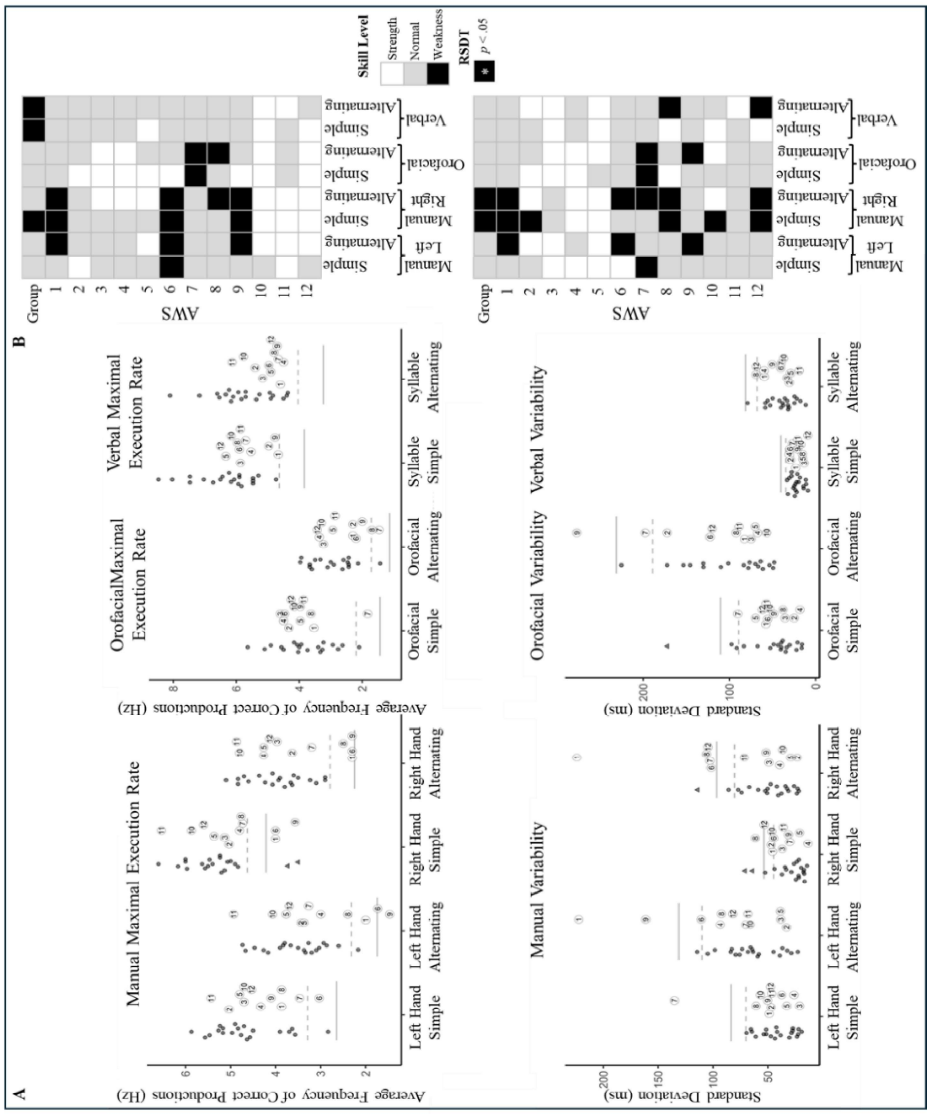
The only weaknesses detected in terms of maximal rate or variability in the execution of speech-related movements concerned AWS 8 and 12 who were significantly more variable than controls in the execution of alternating speech-related movements. This finding, which is in line with AWS 8 and 12's larger variability in the verbal modality of the rhythmic tasks

(see Figure 7), suggests that abnormally variable motor system is associated with stuttering in at least a small proportion of the AWS, but certainly not all.

There were many other group and individual differences. As a group AWS were slower to execute the simple movements with the right hand ($t(15.91) = 2.03$, $p = .03$) and more variable when executing both the simple ($t(19.34) = 2.75$, $p = .006$) and alternating ($t(12.68) = 1.93$, $p = .038$) movements with the right hand. And at the individual level, eight of the AWS showed at least some sign of weakness for at least one measure (maximal execution rate or variability), one hand (right or left) and one condition (simple or alternating). Although these results are interesting, their putative link to stuttering remains unclear as in most of these individuals these weaknesses were not detected for the orofacial movements.

Figure 8

Results of motor assessment



Note. A. Performances of controls (grey filled circles, triangles indicating outliers) and AWS (white circles labeled with their ID) are shown for the two measures of motor abilities, maximal execution rate (left panels) and variability of production (right panels), for the manual (upper panel), the verbal and the orofacial modalities (lower panel). Grey lines

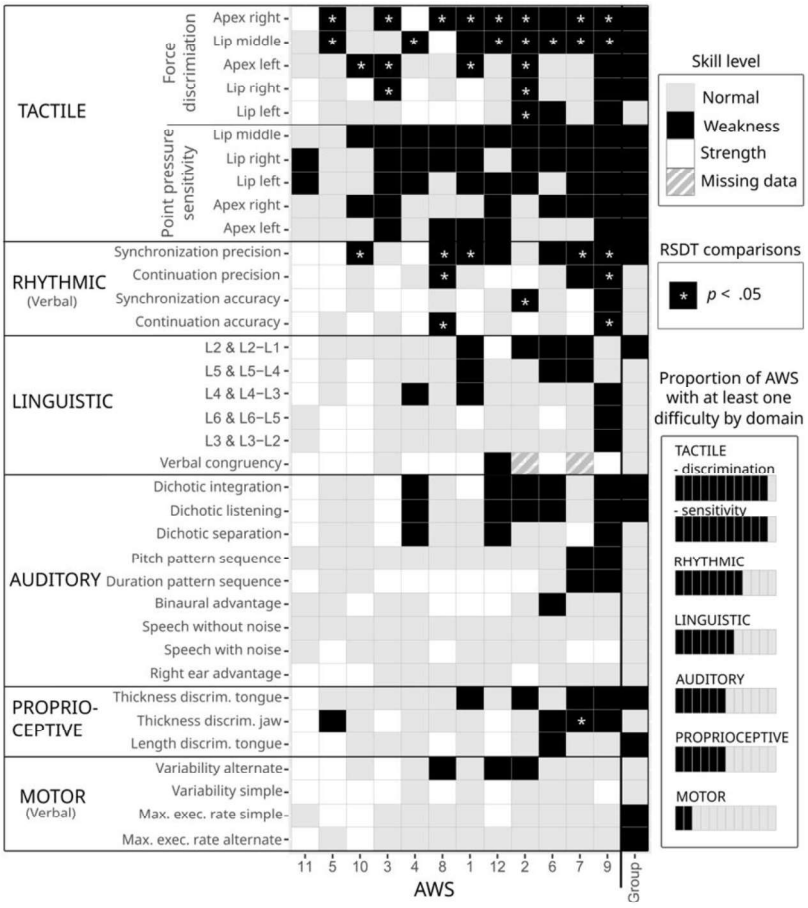
indicate statistical thresholds, corresponding to the first (plain) and fifth (dashed) estimated percentiles of the control distribution. B. Summary of the results. Performances below the fifth percentile threshold are considered as a weakness, performances of AWS numerically above the controls' median indicate a strength. In all other cases, the performance is classified as normal. Cases displaying difficulties in speech related movements (Verbal production in simple and alternating conditions) were compared to their respective performances in control modalities (Orofacial and Manual movements). No significant differences were found in any of the RSDT comparisons (all *ps* > .05).

4.7 Summary of the results

Figure 9 presents the results of all the individual AWS, and the group, in all the speech related tasks and conditions. Three main findings emerged from the data. First, there was large variability across AWS. Some of them showed a weakness in only one domain (AWS 3, 5, 11), while others showed widespread weaknesses spanning all domains under investigation (AWS 7 and 9). Second, some AWS showed above average performance in almost all domains of abilities. AWS 11, for example, performed above average in most of the abilities assessed in the motor, proprioceptive, auditory, linguistic and rhythmic assessments. As discussed in the introduction, this suggests that stuttering is not causally related to a disorder affecting these abilities in these individuals. Third, surprisingly, all the AWS have a tactile perception weakness and, for AWS 3 and 11, this tactile perception weakness appears to be specific and isolated.

Figure 9

Summary of results in all speech related assessments



Note. Performances of each AWS and overall group (horizontal axis) in all speech related assessments (vertical axis) are summarized into three categories: "weakness" (performances below the 5th percentile; black cells), "strength" (performances above the 50th percentile) and "normal" (performances between the 5th and 50th percentile). Domains are sorted from top to bottom according to the proportion of weaknesses observed across AWS. Subjects are sorted from left to right according to the proportion of weaknesses obtained across speech related tasks. RSDT refers to the Revised Standardized Difference test (Crawford & Garthwaite, 2005).

5. Discussion

Sixty years ago, St. Onge & Calvert (1964) asked: "What are we studying when we study stuttering? Whatever it is, is it one, several, or many?" (p.160). The present study aimed to take a first step toward this question by using a multiple single-case approach to chart the strengths and weaknesses of 12 AWS across five domains relevant to speech production. The main question was whether stuttering would emerge as a broadly similar pattern of weaknesses across all individuals, as predicted by domain-general or core-deficit accounts, or as different sets of difficulties across individuals, consistent with the idea of cognitive subtypes. The AWS displayed heterogeneous cognitive profiles. Nevertheless, all exhibited some form(s) of orosensory tactile weakness.

5.1 Cognitive heterogeneity

For long, it is recognized that there is great variability in how stuttering can be expressed in different individuals (e.g., Arndt & Healey, 2001; Blood, 1985; Riley & Riley, 1980; Riley & Riley, 2000; Yaruss et al., 1998). And different attempts have been made to identify distinguishable subtypes of the disorder. However, aside from some exceptions (Alm, 2004; Chang & Guenther, 2020; Civier et al., 2013), previous attempts have focused on different types and severity of symptoms or history of the disorder (see Yairi, 2007 for review). Feinberg, Griffen, & Levery (2000) distinguished chronic tonic from clonic stuttering. McKnight & Cullinan (1987) distinguished a "stuttering-only" group from a group of PWS presenting with associated difficulties. Schwartz & Conture (1988) identified five different groups based on their symptoms, and so on.

Although these classifications are interesting and have enriched our descriptive understanding of stuttering, they are of limited value when the goal is to advance knowledge about the types of dysfunction(s) within the functional architecture supporting fluent speech that may cause stuttering.

Variability in symptom type or severity provides only indirect clues about causation, as symptom-based subtypes may capture different outward manifestations without mapping onto the functional deficits that actually give rise to stuttering. To make progress toward

causal understanding, it is essential to characterize individual differences at the level of the cognitive architecture underlying fluent speech.

In this respect, one striking aspect of the results of this study is the clear demonstration of cognitive heterogeneity in AWS in terms of types and number of weaknesses, normal and strong performance. In almost all the tasks and measures, some AWS were in difficulty – achieving a level of performance corresponding to a weakness – while others performed well above the average performance of the control participants. Of course, caution is warranted in interpreting these findings. Given the large number of comparisons, some deviant scores may well have arisen by chance. At an alpha level of 0.05, approximately 5% of tests would be expected to yield false positives purely due to statistical fluctuation. Conversely, performances within the normal range can also be ambiguous: they might genuinely reflect intact ability, or they could represent Type II errors—failures to detect subtle weaknesses because of variability or other factors.

Despite these considerations, we are confident that the cognitive heterogeneity we report is not merely an artifact of statistical noise but reflects meaningful individual differences in cognitive and sensorimotor functioning among AWS. Across all speech-related tasks, measures, and conditions in the stuttering group (430 total assessments, see Figure 9), 28% were classified as deviant—that is, they yielded p-values below 0.05. This proportion is substantially higher than the 5% false positive rate expected by chance. The probability of observing such a high proportion of deviant outcomes under the null hypothesis is exceedingly low ($p < 10^{-15}$, based on a binomial test). Specifically, for a binomial distribution with 430 trials and a 0.05 probability of success, the critical value at an alpha level of 0.05 is 30, corresponding to roughly 7% of assessments. Therefore, the excess proportion of deviant results (28% observed minus 7% expected) suggests that approximately 21% of the assessments (i.e., 75% of deviant assessments) likely yielded true positives. Importantly, in line with our conclusions, the proportion of deviant performances in the stuttering group significantly exceeded the 5% level expected by chance in all assessed domains, except for the motor domain. Specifically, deviant performance rates in speech-related measures (see Figure 9) were 55% in the tactile domain, 29% in rhythm, 25% in proprioception, 19% in the linguistic domain, 17% in the auditory domain, and 4% in the

motor domain. Binomial tests confirmed that all proportions were significantly greater than chance ($p < 0.0001$), with the exception of the motor domain ($p = 0.7$).

In addition, we believe that performances exceeding the control mean on valid, sensitive tasks raise questions regarding a causal role for the corresponding skill or ability for individuals exhibiting the strong performance. If a theory proposes that stuttering arises from a specific weakness (e.g., an abnormally inaccurate sense of jaw position), then a superior-than-median performance on a well-targeted, sensitive task assessing that ability directly challenges the relevance of that theory for the individual in question.

To be clear, this substantial cognitive heterogeneity among AWS does not imply that there are subtypes of stuttering—let alone that we have identified different subtypes. Our findings reveal variability across a wide range of cognitive domains, but the current study was not designed to determine which of these weaknesses—if any—play a causal role in stuttering. Some of the observed weaknesses are likely to be consequences of stuttering, compensatory adaptations, or even epiphenomenal. Hence, we see our findings as a necessary first step toward the long-term goal of identifying potential subtypes—should they exist—by testing specific causal hypotheses at the individual level.

Regardless of whether distinct cognitive subtypes exist, the heterogeneous cognitive profiles that we report have three main implications that are worth spelling out. The first is theoretical. That at least some AWS were found to have above average performance in all the tasks of the linguistic, inhibitory control, auditory, proprioceptive, rhythmic/timing, and motor assessments indicates that a weakness in the skills and abilities assessed in these evaluations is unlikely at the origin of stuttering in these AWS. Interestingly, cognitive heterogeneity, which naturally leads to population sampling differences across studies, provides a natural explanation for the inconsistent results across studies, which is characteristic of this field of research (St. Onge and Calvert, 1964; Yairi, 2007; Seery et al., 2007; SheikhBahaei, et al., 2023).

A second is clinical. It is probably fair to say that at this juncture, the clinical practice of assessing stuttering is mainly restricted to descriptions of disfluency. There is traditionally little (if any) attempt at exploring the possible causal factors that may contribute to the origin of each patient's symptoms. In this, the evaluation and treatment of stuttering differ from that

of other neurodevelopmental disorders (e.g., dyscalculia, dyslexia). The clinical evaluation of dyslexics, for instance, explicitly aims at discovering the cause of each individual's reading disorder by assessing, in each individual, the set of abilities and skills assumed to be necessary to (learn to) read efficiently. The heterogeneous cognitive difficulties identified in this study call for a similar, more comprehensive cognitive approach to stuttering assessment, as this study provides a clear example of how evaluation can be derived from a specification of underlying skills. Ideally, clinical evaluation of stuttering should include assessments of all the set of abilities and skills assumed to be necessary to support fluent speech to better understand the individual profile of each PWS. Then, similar to the approach taken in the evaluation and treatment of other neurodevelopmental disorders, where individuals receive different treatments that specifically target the putative cause of their symptoms, stuttering treatment should be more personalized. By identifying specific strengths and weaknesses in the cognitive architecture underlying fluent speech, clinicians could tailor interventions to address those specific areas (Sheikhbahaei, et al., 2023). Of course, targeted interventions will not necessarily be effective since a weakness in a particular domain does not necessarily imply that improving this skill through therapeutic intervention will reduce stuttering. However, we argue that, as in other clinical fields, it is the responsibility of the clinician to formulate hypotheses about the potential cognitive contributors to each individual's stuttering and to design interventions that aim at going beyond mere compensation to explore possible causal leverage points. If a targeted intervention proves ineffective, then alternative hypotheses can be generated and other strategies pursued. Such approaches are not intended to replace existing fluency-shaping strategies and cognitive-behavioral interventions, but rather to complement them by engaging with cognitive systems that may plausibly contribute to the persistence or severity of stuttering in a given individual.

A third implication is methodological. Heterogeneity constrains the type of research method that can be useful in stuttering research. Cognitive homogeneity is a prerequisite of group studies aiming at identifying the cognitive, neural correlates and/or genetic dispositions causing stuttering (Caramazza, 1984). The heterogeneity that we report in this study shows that any group study on stuttering clearly violates this assumption. Averaging measures across PWS who differ considerably in terms of the extent and types of weaknesses (and potential mechanisms that underlie their stuttering) is unlikely to be informative. As already discussed in the introduction, collapsing performance arising from individuals with

heterogeneous conditions into one group may lead to inaccurate conclusions, because causal factors of low incidence will be hard to detect, and because group differences may be found although none or very few individuals in the group suffer from an actual weakness. Our findings show that this warning was warranted and indeed provide several examples of these two issues "in action". In Figure 9, for instance, we count 19 instances in which group comparisons did not result in any significant difference between the two groups although up to 33% (4) of the AWS had a weakness in that task. And vice versa, there were several instances where group comparisons resulted in a significant difference between the two groups although very few or none of the AWS actually had a real weakness in the task. In the verbal modality of the motor assessments task, for instance, the two groups differed significantly for both the simple and alternating syllable although no AWS had difficulty in any of the conditions. In sum, for the same reason that we would not seek to discover the neural, genetic or cognitive bases of "a reading disorder" by studying the neural, genetic or cognitive correlates of a mixed group of individuals whose reading disorder is a byproduct of a visual, intellectual, environmental, cerebellar or phonological disorder, we should abandon undifferentiated group comparisons in the study of the causes of stuttering.

Our results also illustrate another peril for the interpretation of structural and functional studies. We have found that in addition to the various patterns of weaknesses detected in tasks assessing abilities that underlie speech production, there are also various associated disorders. Some AWS suffer from tactile weaknesses on the wrist, from a proprioceptive weakness tested on the hand or from difficulty executing fast movements with the fingers. Likewise, previous studies had reported a large set of other associated disorders, such as different types of learning disorders, attention disorders, reading disorders (Blood, Ridenour, Qualls & Hammer, 2003). That some AWS differ from controls along all those lines poses serious difficulties to the interpretation of the results of structural and functional neuroimaging studies. Indeed, in light of our results, it is perhaps not surprising that anatomical and functional differences between AWS and controls have been reported across a very large number of brain regions spanning many subcortical, cortical and cerebellar areas as well as their connectivity (Craig-McQuaide, Akram, Zrinzo & Tripoloti, 2014; Etchell, Civier, Ballard & Snowman, 2018; Matsuhashi, Itahashi, Aoki & Hashimoto, 2023).

Some have argued that a way forward could be to study subtypes of PWS based on their symptoms or history (Yairi, 2007; Yairi & Ambrose, 2013; Alm, 2005). While intuitively appealing, this strategy can only be effective if there is a systematic, one-to-one relationship between external symptoms—or developmental trajectories—and underlying cognitive mechanisms. But this is unlikely to be the case. Similar symptoms may well result from impairment to different cognitive components underlying speech (e.g., Chang & Guenther, 2020). In addition, PWS who suffer from similar symptoms may well suffer from different constellations of different associated disorders. And vice versa, individuals who present with different symptoms or history may well present with very similar cognitive profiles. The way stuttering manifests itself, through avoidance, blocks, repetitions or escape behaviors may well depend on a combination of personality traits and learned behaviors rather than reveal different causal mechanisms in action.

To make real progress, we believe subtypes—if they exist—should not be defined a priori based on symptoms or clinical presentation, but should instead emerge from careful analysis of the cognitive precursors and mechanisms of fluent speech. In our view, rather than pursuing more group-based studies, future research should prioritize fine-grained examination of cognitive profiles and potential contributors to stuttering at the individual level, followed by testing of causal hypotheses through targeted single-case intervention or training studies. Carefully tailored interventions indeed offer a powerful method for testing causal hypotheses. If a particular skill or ability contributes causally to stuttering, then strengthening that skill or ability should lead to measurable decrease in stuttering and improvements in fluency (Danks & Davis, 2023 for discussion).

Of course, this approach is difficult, and likely impossible in PWS who suffer from many different weaknesses and difficulties. Take AWS 9 from our study, for instance, who suffers from weaknesses in almost all the domains. It would be hard to know what to infer from this profile. However, our hope is that this approach will also allow us to identify some PWS with relatively "pure" disorders within the speech production architecture, allowing us to generate clear and testable hypotheses. Here, for instance, we report that all of the AWS that we tested suffered from an orosensory tactile disorder. And in two of them (AWS 3 and 11), this orosensory tactile disorder was the only weakness that we observed. This finding invites a clear and tractable hypothesis about a possible causal mechanism of stuttering, as

we discuss below, that invites further intervention studies to test the existence of a causal relationship between these disorders. Likewise, other studies could identify other PWS who suffer from a selective disorder within the linguistic system, or in timing, or in rhythm, or in motor abilities and so on. Although this kind of individualized, mechanism-focused approach may be unfamiliar in the stuttering literature, it has started to yield valuable insights in related fields. Case studies examining individual cognitive profiles have helped advance our understanding of the functional underpinnings of conditions such as dyslexia, dyscalculia, developmental language disorder, and aphasia (De Visscher & Noël, 2013; Thompson, 2007; Träff et al., 2017; Ramus et al., 2003; White et al., 2006) —even if such studies remain relatively uncommon, and most research in these areas continues to rely primarily on group-comparisons.

5.2 A possible contribution of a tactile orosensory disorder?

All the AWS that we tested showed from some form(s) of orofacial tactile weakness. The higher incidence of orofacial tactile weakness than any other types of weakness (linguistic, inhibitory, auditory, proprioceptive, rhythmic, motor) cannot be accounted for by the hypothesis that these tests were more sensitive. If anything, the orosensory tactile tasks were the least sensitive as they were characterized by a ceiling effect in the control participants. This higher incidence cannot be explained either by the total number of tests assessing this ability. As noted in the previous section of the discussion, for all speech-related tasks, measures and conditions being considered within each domain, the proportion of deviant performances (i.e., assessments with p-values < 0.05) was markedly higher in the tactile domain (55%, 95% binomial proportion confidence interval [46%, 64%]) compared to the other domains. The next highest proportion was observed in the rhythm domain, with 29% deviant performances (95% CI [17%, 44%]). A one-sided Fisher's exact test confirmed that the proportion in the tactile domain was significantly greater than in the rhythm domain ($p = 0.02$).

This finding is rather surprising. Previous data on the orosensory tactile ability of PWS is scarce and somewhat conflicting (see pp.266-267 in Bloodstein et al., 2021 for a recent overview). The available data comes from three types of tasks. The first is the oral

stereognosis task, in which participants are required to discriminate or recognize shapes placed in their mouth. Using variations of this task, several studies have reported evidence for a group difference between PWS and normo-fluent controls (Zsilavecz, 1981; Gunderson, 1972; Martin et al., 1981; Stewart et al., 1985) and others have failed to report any difference (Cook et al., 2011; Jensen et al., 1975; Hutchinson & Ringel, 1975; Carpenter & Sommers, 1987). These latter findings are not incompatible with ours. Although it is not totally clear what this task evaluates, stereognosis probably depends more on proprioceptive abilities and is unlikely to be affected by the type of subtle orosensory tactile disorders found in our study. Indeed, the mixed findings in this literature may find a natural explanation in the sampling bias that should be expected if proprioceptive difficulties affect only a small proportion of the PWS, like our results suggest. Two studies reported intact two-point discrimination tested on both the lips and tongue of PWS, suggesting intact spatial acuity (Jensen et al., 1975, Zsilavecz, 1981). Finally, in line with the results reported herein, several studies have reported increased sensitivity thresholds to vibrotactile stimuli (Petrosino et al., 1987; Fucci et al., 1991) and decreased tactile vibration acuity (Fucci et al., 1991, Harris, Fucci, & Petrosino, 1991) in PWS.

A limitation of these studies is that it remained unclear whether the reported orosensory weaknesses were specific or rather appeared in the context of concomitant deficits affecting other cognitive abilities underlying either performance in this type of task (e.g., attention) or speech production (e.g., rhythmic, linguistic, auditory, and so on). In this respect, the results reported herein allow clearer conclusions. All the AWS (11/12) who presented with a weakness in orosensory tactile force discrimination performed also significantly worse when tested on an orofacial area than on the wrist. And 9 of the 11 AWS who presented with a weakness in point pressure sensitivity for at least one orofacial area had normal performance in a control condition. These disproportionate weaknesses measured on orofacial versus control body parts imply that these difficulties are not a mere consequence of task-related domain general difficulties (e.g., attention). And indeed, most of the AWS included in this study performed at a rather good level of performance in all the other tasks and domains. AWS 3 and 11 in particular did not show any weakness in any other domain.

Of course, this does not imply that the AWS included in our study do not have any other disorders. Although the set of abilities that we tested is arguably rather large, it does not

provide an exhaustive inventory of all the speech-related processes and sub-processes that might contribute to stuttering. Several speech-related processes and abilities, including some that have been proposed to contribute to stuttering were not tested. For instance, although several theories suggest that audio-motor integration may be a source of difficulty at the origin of stuttering (see Bradshaw et al., 2021 for review), methodological limitations made it impossible for us to include an experiment appropriate to test this hypothesis in our battery of tasks (e.g., compensation and adaptation to experimental modifications of the auditory feedback). Likewise, multi-sensory integration is widely assumed to play a critical role in speech motor control (Hickok, et al., 2011; Guenther, 2016) – it allows using feedback from different sensory modalities for the on-line monitoring and potential correcting of speech errors. Some empirical evidence suggests that some PWS may suffer from a difficulty in accurately synchronizing somatosensory and auditory information. The study of somatosensory-auditory integration is typically investigated in tasks involving finger tapping to a metronome, where participants are instructed to tap their finger on a surface along with a rhythmic train of tones. In such "paced tapping tasks", PWS have been reported to show a greater "Negative Mean Asynchrony" (NMA) (Sares, Deroche, Shiller & Gracco, 2019; Falk, Muller & Dalla Bella, 2015; Slis et al., 2023). We had initially hoped to measure NMA from the data collected in our rhythmic synchronization and continuation task. Unfortunately, after a careful check of the timing of the (first) sound emitted by our experimental software, it was judged too unreliable to calculate this index. We also omitted an evaluation of participants' lip proprioceptive abilities. We did so because the only study assessing lip proprioceptive ability in AWS so far (De Nil & Abbs, 1991) reported that lip difficulties (detected in 2/6 participants) were less prevalent than and always accompanied by tongue (3/6) and jaw (4/6) deficits. And although we tried to select sensitive tasks and valid measures, it could also be that, in some cases, the tasks that we used to assess some core hypotheses did not target with enough precision the type of linguistic, auditory, motor, rhythmic or proprioceptive abilities that are in fact impaired in PWS. It could well have been, for instance, that the AWS that we tested would have had difficulty in a movement rate task with more difficult movements, or with longer sequences, or if we had tested the variability of motor abilities under cognitive load. It could also well be that a more complex tapping task would have facilitated to the detection of rhythmic difficulties (e.g., Webster, 1990). And so on.

Nevertheless, the finding that an orosensory tactile disorder was both necessary (present in all) and sufficient (the only weakness detected in two AWS) for stuttering to occur in the

small sample of AWS that we tested in this study suggests a stronger association between stuttering and orosensory tactile disorders than previously assumed.

But before going farther, and in order to avoid misunderstandings, it may be useful to clarify that the present findings and previous literature together already largely constrain the space of viable hypotheses about the severity of orosensory disorders that may or may not cause stuttering. Past studies examining the role of orofacial somatosensory information in speech control have documented that congenital (MacNeilage, Rootes, & Chase, 1967), acquired (Hoole, 1987) and experimentally induced (Putnam & Ringel, 1976; Ringel & Steer, 1963; Scott & Ringel, 1971) loss of somatosensation in the oral cavity led to imprecision in speech articulation but not to stuttering. This suggests that stuttering may be caused by subtle – not severe – orosensory tactile disorders, in line with the subtle tactile disorders that we found in the AWS tested in this study.

At first glance, this idea may seem counterintuitive. However, non-monotonic relationships between disorder severity and symptoms are well documented across sensory systems. Although mild binocular imbalance produces disturbing diplopia, in case of more severe imbalance the brain suppresses input from one eye, eliminating double vision. Partial vestibular loss produces vertigo, but complete bilateral loss abolishes dizziness because the unreliable channel is no longer integrated. In speech motor control, small delays of auditory feedback disrupt fluency more strongly than longer ones or a complete absence of feedback (e.g., Boller et al., 1978; Howell & Archer, 1984; Stuart et al., 2002). These cases illustrate the principle of precision weighting in predictive coding and Bayesian cue integration frameworks (e.g., Ernst & Bülthoff, 2004; Körding & Wolpert, 2004), whereby highly unreliable inputs are down-weighted or ignored, while moderately unreliable inputs remain integrated and are therefore more disruptive. By analogy, we propose that subtle orosensory impairments may destabilize speech motor control and manifest as stuttering, whereas severe impairments may lead the system to down-weight or ignore orosensory input, resulting in imprecise articulation without disfluency.

Future research will be needed to test whether, and if so how, subtle orofacial tactile disorders indeed contribute causally to stuttering. Alternatives include the possibility that these weaknesses are a consequence of reduced or qualitatively different speech experience in PWS or that they are a mere epiphenomenon co-occurring systematically with stuttering

without being causally related to its onset or persistence. The hypothesis of a causal role appears plausible, though, and potentially illuminating to account for a vast array of characteristics of stuttering currently in need of a common explanatory framework. Hence, although our current findings do not establish a causal link, they motivate a hypothesis that warrants further investigation.

Somatosensory information has long been known to be important for speech production (Ringel & Steer, 1963; Scott & Ringel, 1971; Tremblay, et al., 2003; see Guenther, 2016; Kent et al., 2024 and Parrell & Houde, 2019 for recent reviews) and the hypothesis that stuttering could be rooted in deficient orosensory processing is far from new (Mysak, 1960; Van Riper, 1971). According to Van Riper (1971), for instance, stuttering would emerge in children who fail to transfer from auditory to somatosensory control of speech during development (see also Lane & Tranel, 1971).

Within the DIVA model of speech motor control, three main hypotheses have been proposed to account for stuttering (Max et al., 2004; Civier et al 2010, 2013): (1) «Unstable or Insufficiently Activated Internal Models» (Max et al., 2004), leading to abnormal detection of speech production errors and causing a frequent triggering of motor-reset correction mechanisms (2) «Weak Feedforward Control and Overreliance on Afferent Feedback» (Max et al., 2004, Civier et al., 2010) causing control instabilities and similarly frequent triggering of motor-reset correction mechanisms, and (3) temporal sequencing failures during syllable selection and initiation (Civier et al., 2013). These hypotheses suggest several different causal pathways by which subtle orofacial tactile disorders—that is, degraded but still integrated tactile input— may lead to stuttering. Interestingly, these different causal pathways may account for different symptoms, characteristics of stuttering, and for the effect of different fluency-enhancing conditions.

In the first perspective, unstable or insufficiently activated internal models would lead to an abnormal detection of speech production errors, causing an abnormal triggering of motor-reset correction mechanisms, even when no errors exist, resulting in a repetition (Max et al., 2004; but see Postma, 1992 for a critical view on this idea). We see at least two ways by which a mildly disrupted somatosensory tactile system could result in such scenario. The first is by hampering motor learning, leading to less precise, possibly more unstable internal models (Darainy, Manning & Ostry, 2023; Ebrahimi & Ostry, 2024). Second, given the role

of orofacial tactile cues in providing spatial and positional information, a subtle orofacial tactile disorder would likely result in a combination of inaccurate estimations of the initial configuration of the articulators, leading to inaccurate motor commands, and inaccurate feedback signal and correction, triggering a motor reset in which the system attempts to repair the error by restarting the syllable.

The second perspective suggests that weak feedforward control, compounded by an overreliance on afferent feedback, would create motor control instabilities and similarly abnormal triggering of motor-reset correction mechanisms. This scenario could also result from a deficient somatosensory tactile system as repeated errors and noise in the estimation of the configuration of speech articulators would likely interfere with the acquisition and update of feedforward models during motor learning. A weakened, less efficient and reliable feedforward system would in turn result in an increased reliance on sensory feedback (auditory and/or somatosensory) heightening motor control instabilities and the likelihood of speech errors and motor resets. Civier et al. (2010) demonstrated this idea through computer simulations, though they focused on auditory feedback alone. They noted, however, that "in addition to auditory feedback control, the biasing may be toward somatosensory feedback control as well" (Civier et al., 2010, p.248).

Additionally, a deficient somatosensory tactile system may further increase the reliance on auditory feedback (Patri et al. 2019), further increasing motor control instabilities, and potentially stuttering, due to the longer processing delays inherent to auditory feedback (e.g., Houde & Nagarajan, 2011; Parrel & Houde, 2019; Postma, 2000) and to the discontinuous nature of this sensory modality as compared to somatosensation. In line with this possibility, AWS have been reported to be less affected by manipulation of the somatosensory feedback than normofluent controls (Loucks, & De Nil, 2006a) and simulations in the DIVA framework have shown that increasing reliance on auditory feedback replicates multiple features of stuttered speech (Civier et al., 2010; Civier et al., 2013). This may explain why alterations of the auditory feedback - by inducing delays, by masking or by inducing shifts in fundamental frequency - reduces stuttering frequency (Bloodstein & Ratner, 2008; Foundas et al., 2013; Kalinowski et al., 1993). At the same time, it is important to acknowledge a key issue for this hypothesis: if increased reliance on auditory feedback were the primary mechanism, PWS should exhibit larger compensatory

responses to unexpected auditory perturbations. Yet, current evidence does not support this prediction (see Bradshaw et al., 2021).

Furthermore, a deficient somatosensory tactile system could at least in part explain why slowing speech improves fluency in PWS (Blomgren & Goberman, 2008; Logan et al., 2002). Speech rate has different potential impacts on articulation such that, depending on the speaker, context, and articulator, increasing speech rate may be achieved by a combination of an increase in velocity, by a decrease in movement amplitude (articulatory undershoot), and increased coarticulation (see Berry, 2011 for a review). Any of these articulatory adaptations will likely decrease the saliency of the somatosensory cues available for speech motor control, hence further degrading state estimation, motor planning and motor coordination. Slower speech, by contrast, would increase the precision and saliency of somatosensory signals, by allowing, in particular, longer sensory accumulation and more intense contacts between speech articulators.

The third perspective, involving temporal sequencing failures, posits that stuttering may stem from disruptions in the timing of syllable transitions. Since somatosensory tactile information is essential for the continuous monitoring of articulatory positions before initiating speech sounds, it likely provides crucial information for the proper triggering of the initial sounds of utterances (Loucks et al., 2010). Therefore, a failure to confirm correct articulator positions for initiating voicing when the motor program is ready, would likely result in blocks on the initial sound of utterances – a common difficulty for PWS. Similarly, orosensory tactile impairments may lead to timing delays during connected speech. In the GODIVA model (Bohland et al., 2010), stuttering arises from delays in the deactivation of the ongoing syllable (leading to prolongations) or in the activation of the subsequent syllable (leading to blocks). Among the mechanisms proposed to control this temporal sequencing, Civier et al. (2013) considered that an efference copy of the current motor command is used to estimate and anticipate the imminent completion of the current syllable to trigger the transition to the next. While Civier et al. (2013) suggested that this anticipation relies solely on efference copy, it is reasonable to assume that this prediction would actually involve the integration of both efference copy and sensory feedback, as currently implemented in state feedback control models (Houde & Nagarajan, 2011). In this scenario, various types of orosensory tactile impairments could impact state estimation and disrupt the timing of

syllable transitions. For example, noisier orosensory information could increase state estimation uncertainty, inducing temporal delays and instabilities. Longer transmission or processing delays of orosensory information have actually been suggested as a possible explanation for the greater anticipation of taps in people who stutter during synchronization tasks involving tapping along to an auditory metronome (Slis et al., 2022). Furthermore, this perspective could also explain why delayed auditory feedback improves fluency in PWS but disrupts fluent speakers. If stuttering stems from misaligned sensory feedback, delaying auditory input may help realign it with orosensory feedback, reducing fluency disruptions. Delaying auditory feedback in fluent speakers, in contrast, would disrupt speech by causing a misalignment between the two sensory modalities.

In sum, through these three combined hypotheses, an orosensory tactile origin of stuttering offers a potential unified account of several key characteristic of the disorder, such as its developmental onset (Van Riper, 1971), its characteristic difficulty with initial sounds, increased articulatory errors, and the finding that they seem less affected by manipulation of the somatosensory feedback than normofluent controls (Loucks, & De Nil, 2006a), but aided by alterations of the auditory feedback by inducing delays, by masking or by inducing shifts in frequency. This hypothesis also offers interesting insights into different fluency-enhancing techniques, such as external timing cues (e.g., metronomes, choral speech, singing, Andrews et al., 1982; Falk et al., 2016; Frankford et al., 2022; Kiefte, & Armson, 2008; Wiltshire et al., 2024), by suggesting that these cue may compensate for crucial but unreliable somatosensory input, and the use of unfamiliar speaking modes, which could be explained by the fact that these modes do not yet involve learned somatosensory targets such that impairments in somatosensory feedback would have no consequences on the detection of speech motor errors. As speakers adapt to new modes, however, somatosensory targets are formed, and stuttering would tend to re-emerge.

6. Conclusions

In sum, this study leads to three main conclusions. First, our findings reveal clear cognitive heterogeneity among adults who stutter (AWS) across a range of cognitive and sensorimotor domains previously hypothesized to contribute to the disorder. Although it remains unclear whether—and if so, which—of these factors contribute causally (see below),

this heterogeneity underscores the need for more individualized approaches in both research and intervention.

Second, all twelve AWS in our sample exhibited an orosensory tactile disorder. While this finding does not imply that stuttering is always caused by orosensory impairment, nor that all relevant contributors have been identified, it does suggest that tactile orosensory function may play a more central role in stuttering than previously recognized. As such, this domain represents a promising target for future theoretical and empirical investigation.

Third, some AWS in our study demonstrated above-average performance relative to controls across multiple domains, including psycholinguistic access, rhythmic and auditory processing (e.g., dichotic listening), proprioception, and speech motor control (in terms of speed, accuracy, and consistency in both simple and complex tasks). This observation has important theoretical implications. It suggests that weaknesses in these specific skills and abilities may not be necessary contributors to stuttering. In other words, stuttering in at least some individuals may not be due to disorders in the types of linguistic, rhythmic, or motor abilities tested here. This places meaningful constraints on explanatory models of stuttering by identifying which types of impairments cannot be considered universal. To avoid misunderstanding, we emphasize that this conclusion applies strictly to the specific skills and abilities we tested. Our battery does not capture all possible aspects or subcomponents of these domains. For example, stuttering may still plausibly arise from difficulty coordinating speech under cognitive load or from irregularities in planning speed—possibilities not directly addressed by the present tasks. Theoretical models proposing causal roles for these domains must now clarify how their predictions accommodate individuals with such profiles. In this way, our findings help narrow the range of viable hypotheses and provide a clearer framework for future research.

We would like to conclude by clarifying six important aspects regarding the scope of the present study. First, the study is unlikely to offer a representative sample of the adult stuttering population at large. To our knowledge, there is currently no scientifically agreed-upon standard for defining an "appropriate" participant in studies of developmental stuttering—whether in terms of co-occurring conditions, symptom severity, or age of onset. In designing this study, we followed common practice in the literature by including individuals who (1) had received formal diagnoses, (2) reported stuttering-related difficulties

in daily life, and (3) self-identified as people who stutter at the time of testing. These criteria, in our view, provide a strong basis for inclusion in a study focused on the cognitive underpinnings of stuttering. It is clear, however, that this sample does not capture the full demographic, clinical, or cognitive diversity of AWS. Future research with larger samples will be essential to progress in this direction.

Second, this study was not designed to estimate the prevalence or co-occurrence of specific disorders within the broader population of AWS. Our objective was to explore the range and nature of cognitive variability across domains that have been theoretically linked to stuttering, not to make epidemiological claims. While our findings reveal the presence of orosensory tactile disorders and other patterns of individual variability, they do not allow us to determine how common these disorders—or combinations thereof—are across the stuttering population at large. Accurately addressing such questions would require large-scale, representative studies with appropriate sampling strategies. Instead, the present study aims to identify candidate mechanisms and illustrate the cognitive heterogeneity that exists in a small sample of AWS, thereby laying the groundwork for future research that can assess incidence patterns more systematically.

Third, our study does not offer an exhaustive inventory of the cognitive and sensorimotor abilities involved in fluent speech. We selected a battery of tasks designed to probe skills and abilities frequently discussed in the stuttering literature. But any such battery necessarily reflects methodological choices and exclusions. For example, we did not include tasks assessing audio-motor integration (e.g., adaptation to altered auditory feedback), nor did we examine multisensory integration, despite their theoretical relevance. We hope that future work—ideally involving larger samples and more comprehensive batteries—will build on our approach to further map the cognitive architecture underlying fluent speech. Still, these omissions do not undermine the validity of the three core conclusions outlined above.

Fourth, this study does not establish causality. The observed disorders cannot be assumed to cause stuttering; they may instead reflect consequences, compensatory mechanisms, or unrelated comorbidities. Interventional, longitudinal, and mechanistic studies are needed to move beyond correlation and address causal inference directly.

Fifth, as a corollary, our study does not provide evidence for distinct cognitive subtypes of stuttering—that is, subgroups of AWS whose stuttering arises from different underlying dysfunctions within the speech production system. While the heterogeneity we observed may eventually help define such subgroups, doing so requires demonstrating a causal relationship between specific disorders and the onset or maintenance of stuttering. Our contribution lies in documenting systematic individual variability across theoretically relevant domains—a necessary first step toward that longer-term goal and also in suggesting caution against assumptions of homogeneity in group-level research.

Sixth, our study focuses explicitly on *proximal* contributors to stuttering—that is, factors operating within the functional architecture that directly supports fluent speech. We did not examine *distal* contributors such as genetic variation or structural neuroanatomical anomalies, nor *modulatory* influences like anxiety or speech rate. The latter are undoubtedly important and complementary research avenues. However, in our view, future work must first advance in two key directions before such approaches—particularly genetic and neuroimaging studies—can fully realize their potential: (1) determining whether the proximal cognitive contributors to stuttering are homogeneous or heterogeneous across individuals, and (2) developing a clearer understanding of the co-morbidity between stuttering and other cognitive disorders, including its variability. We firmly believe that without progress on these two fronts, it will remain difficult to interpret findings from group-based neuroimaging or genetic studies—approaches that typically assume causal homogeneity and overlook potential (heterogeneous) co-morbidity within the stuttering population.

Taken together, our findings highlight the value of individual-level analyses in studying stuttering. By moving beyond group averages, future research can better test specific causal hypotheses and, in doing so, lay the groundwork for more personalized, mechanism-based clinical interventions.

7. Acknowledgement.

We are deeply grateful to Julie Viardot et Cécilia Caravaca for their invaluable help with participants recruitment and data collection for this study.

8. Supplementary materials

The supplementary materials can be found in the appendix of this thesis and include all figures representing analyses conducted without excluding outlier participants. You will also find a section presenting a complementary study, carried out in our laboratory by Jean-François Patri, which demonstrates some performance consistency in tactile difficulties for several of the AWS participants involved in the study described in this chapter.

General discussion

At the beginning of Chapter 1, we posed the following question: *What failures in the processes that govern fluent speech give rise to the broad manifestations observed in stuttering?* As outlined in the introduction, despite decades of extensive research, our understanding of the cognitive underpinnings of stuttering remains limited. Although the variability in the manifestation of stuttering has been recognized since the 19th century (see Ssikorski, 1889), suggesting the presence of potentially diverse underlying cognitive mechanisms, this recognition has had little influence on how stuttering is currently investigated. Yet, even the slightest possibility that stuttering may arise from multiple distinct mechanisms carries significant clinical and research implications.

In the present thesis, we proposed an alternative framework for studying stuttering, with the goal of directly demonstrating empirical evidence of its cognitive heterogeneity and reflecting on potential causal contributors. We believe that this approach offers a step forward in the effort to finally *unty the knot of stuttering*.

1. Summary of the results

Despite the widespread recognition that PWS display distinct patterns of strengths and weaknesses across the abilities that support fluent speech (see Yairi, 2007), this acknowledged diversity has had limited influence on research methodology. Stuttering continues to be treated as a "pathognomonic monolith" (St. Onge & Calvert, 1964, p. 160), reflecting an implicit assumption of causal homogeneity. This assumption has contributed to the predominant use of group-level analysis designs in the field.

In Chapter 1, we discussed and critically examined the limitations of such designs. We argued that group-level analysis rest on two fundamental assumptions: (1) a high degree of homogeneity within the group being studied, and (2) consequently, that differences between groups can be meaningfully captured by comparing average performance. While comparing average performance can yield informative results under the first assumption, its validity is significantly weakened in the presence of heterogeneous groups. In such cases, comparisons based on average performance may produce population sampling differences across studies, contributing to mixed and often inconsistent findings in the literature. Moreover, they can mask or overestimate potential underlying mechanisms.

Overall, group studies have suggested that people who stutter (PWS) exhibit difficulties across nearly all cognitive domains involved in fluent speech, thereby portraying stuttering as a complex and diffuse disorder with limited potential for targeted intervention. However, in the field of stuttering, where heterogeneity is not only prevalent but arguably central, the methodological highlighted in this thesis are particularly salient and underscore the need for alternative research approaches. We advocate for the use of single-case designs, which can offer the level of resolution necessary to directly demonstrate empirical evidence of cognitive heterogeneity and to identify potential causal contributors.

In Chapter 3, we combined group-level and single-case analyses to investigate whether AWS exhibit phonological working memory difficulties independent of overt speech production and/or broader working memory weaknesses. At the group level, AWS tended to show phonological working memory difficulties within the context of general working memory impairment. However, distinct performance patterns emerged at the individual level. These findings underscore the importance of individual-level analysis and broad cognitive

assessments to better understand the possibly heterogeneous mechanisms that underly stuttering. Moreover, they challenge some theoretical accounts.

In Chapter 4, the individual performance of twelve AWS was assessed using an extensive battery of tasks, with the aim of establishing individual cognitive profiles and identifying potential contributing factors to stuttering. The results revealed heterogeneous cognitive profiles across participants. However, they all exhibited an orosensory tactile deficit. For some individuals, this tactile deficit was the only observable difficulty, suggesting that orosensory processing has not received sufficient attention in past research regarding its potential link to stuttering.

While some other analyses could have been conducted, we deliberately chose to omit them for reasons detailed in the next section,

For example, past research has frequently investigated whether there is a correlation between the degree of stuttering severity and empirical data collected. Such correlations are often viewed as indicators of possible causal relationships. However, these analyses are meaningful only if the severity measures used are reliable. One commonly used tool for measuring stuttering severity is the Stuttering Severity Instrument–Fourth Edition (SSI-4) developed by Riley (2009). The SSI-4 assesses severity by combining frequency, duration, and physical concomitants of stuttering into a composite score. While commonly used for clinical and research purposes, the SSI-4 captures the severity of a PWS only at a specific point in time and in a particular context. Furthermore, studies such as Constantino and collaborators (2016) have shown that stuttering severity scores assessed with the SSI-4 can vary significantly over time within the same individual. In some instances, fluctuations in stuttering severity are so pronounced that an individual who stutters' score may fall below the threshold typically used to classify someone as a stutterer. This variability highlights limitations of the SSI-4 in representing the severity consistently across different moments and conditions. Therefore, the inherent variability and context-dependence of severity measures mean that such analyses should be interpreted with caution. These considerations motivated the choice not to pursue such analyses in this thesis.

It would have been interesting to examine whether the specific loci of tactile difficulties observed in each AWS were associated with the loci of their speech disfluencies.

Specifically, we could have investigated whether individuals who stuttered more frequently on certain sounds also exhibited tactile deficits in the articulators involved in producing those particular sounds. Such an analysis could have clarified potential links between orosensory processing and the behavioral characteristics of stuttering for each participant. However, several challenges made it difficult to pursue this line of analysis. First, all our participants exhibited only very mild to mild stuttering at the time of testing, which limited the number of observable stuttering events and thus the robustness of any interpretation. Second, it was often difficult to identify precisely which sounds triggered each stuttering episode. Third, many participants presented tactile deficits specifically involving the corners of the mouth, an area engaged in the articulation of a wide array of sounds. This anatomical overlap made it impossible to interpret meaningful associations unique to each participant.

Finally, we could have run exploratory cluster analysis. Although exploratory cluster analysis can occasionally identify meaningful subgroups, we opted not to employ this technique in the present study. Cluster analysis requires a sufficiently large sample size to yield statistically reliable and interpretable results (see Dalmaijer, Nord, & Astle, 2022), as well as some theoretical justification or hypotheses to inform cluster characterization. Our dataset consisted of only twelve participants, and we lacked clear expectations about group structure or the necessary background knowledge to guide such analysis. Under these conditions, any clustering solution would be unreliable, unlikely to yield robust or generalizable insights, and could even lead to misinterpretation. Therefore, we determined that cluster analysis was not appropriate for our data.

Chapters 3 and 4 further emphasize the cognitive heterogeneity observed among PWS. We will discuss the implications of this interesting finding in the next section.

2. Cognitive heterogeneity and its implications for the search of causation

One of the main findings of the present thesis is the direct empirical demonstration of cognitive heterogeneity. While some may consider this finding insignificant, suggesting that cognitive heterogeneity could already be inferred from the mixed and inconsistent group-level results reported in previous studies, we argue that this perspective is mistaken.

Firstly, mixed and inconsistent results across group studies do not reflect true cognitive heterogeneity. Such variability can arise from numerous methodological sources, including standard practices in experimental design and data exclusion (Siritzky et al., 2023), insufficient statistical power, and limited task sensitivity. Secondly, variability across group-level studies indicates that two distinct samples of PWS perform, on average, differently compared to two samples of PWNS on tasks designed to assess the same ability, thus leading to inconsistencies in the findings. However, this does not provide clearer insight into the individual cognitive profiles of the PWS investigated, nor does it demonstrate cognitive heterogeneity directly. Finally, although it is widely acknowledged that stuttering is characterized by substantial variability among individuals, research methodologies have largely continued to treat stuttering as a homogeneous disorder at the cognitive level (Yairi, 2007).

We are therefore confident that the evidence presented in this thesis offers a novel and complementary contribution that was previously unavailable: PWS demonstrate clear cognitive heterogeneity in the types and numbers of difficulties and strengths they exhibit regarding abilities that are necessary to speak fluently. Some may argue that we may not have detected difficulties in certain domains (e.g., motor abilities) because our tasks lacked sufficient sensitivity, or that certain measures (e.g., some somatosensory abilities) were not sufficiently reliable due to ceiling effects. This is a valid concern, as the sensitivity and reliability of cognitive tasks are crucial for providing meaningful insights into the relationship between the cognitive processes being measured and the disorder under investigation (Zorowitz & Niv, 2023).

We wish to address this point with two arguments that support the adequacy of our chosen tasks. First, for most of the abilities assessed, the distribution of performance in the control group was normal, suggesting that these tasks were neither too easy nor too difficult and were able to capture meaningful variability. Although some measures of orofacial tactile perception exhibited (near) ceiling effects, reflecting limitations in available measurement tools, we still observed that the performance of all AWS on (some of) these tasks diverged substantially from the control group. Furthermore, in cases where significant deficits were found in orofacial areas, all the AWS demonstrated at least one poorer performance not only relative to controls, but also in comparison to a corresponding control body part. This pattern

strongly suggests that the observed difficulties reflect specific orofacial tactile impairments rather than nonspecific factors such as attention or task comprehension.

Second, regarding the difficulty level of certain tasks, we contend that task complexity should serve as a criterion only when there is theoretical justification to assume that the disorder under investigation manifests primarily under such demanding conditions. In the present study, all tasks were selected based on well-established theoretical frameworks, ensuring that their complexity was appropriate and relevant to the mechanisms believed to underlie the disorder.

For instance, a reviewer suggested that the absence of motor difficulties in our participants resulted from the use of "a rudimentary diadochokinetic tasks (DDK)", which they considered a poor measure of motor ability. However, we respectfully disagree. The motor tasks employed in this thesis were specifically designed to test a central prediction of the Speech Motor Skill (SMS) hypothesis (Namasivayam & Van Lieshout, 2011; see Section 3.1.5 of Chapter 1): that conditions destabilizing movement synchronization (e.g., increased speech rate) exacerbate stuttering by imposing greater demands on motor control. If this prediction was correct, stuttering-like behaviors would have emerged, especially in the fast alternating condition. We might also have observed increased inter-movement variability, suggesting destabilization of the motor system. Additionally, we could have acknowledged a significant decrease in speech rate as a compensatory mechanism; however, this was not observed. The choice to use DDK tasks aligns with prior studies that have empirically tested the SMS hypothesis (Namasivayam & Van Lieshout, 2008; see also Van Lieshout, 2017). Moreover, DDK tasks are known to place maximal demands on articulatory speed and accuracy, as suggested by the SMS framework, making them highly suitable for revealing potential motor limitations (see Kent, Kim, & Chen, 2022). Additionally, the DDK paradigm is advantageous in that it isolates motor skill from psycholinguistic influences, as the task is independent of higher-level language processing.

This being said, we wish to emphasize that demonstrating cognitive heterogeneity is not synonymous with identifying the cause(s) of stuttering, nor does it allow us to delineate distinct subtypes of stuttering, should they exist. We are fully aware that our results do not provide such evidence. Nevertheless, we believe that our approach represents a necessary

step toward these goals. Here, we discuss why this step is crucial and consider the implications for causation.

2.1 The importance of setting a logical framework

As early as 1998, Attanasio and collaborators observed that the search for causal mechanisms in stuttering "has been hampered by the inappropriate use of the representativeness heuristic, a method of reasoning which proposes that the effects of a disorder reflect its cause" (p. 265). They argued that this heuristic has led to the problematic assumption that the observable features of stuttering necessarily mirror its underlying causes. In other words, the presence of difficulties across nearly all cognitive domains involved in fluent speech, as well as the widespread structural and functional brain differences among PWS—as revealed by group-based studies—has fostered the belief that stuttering must itself be explained by complex, multifactorial mechanisms.

However, this line of reasoning is not necessarily valid. Simple biological or cognitive factors can, in fact, give rise to surprisingly complex behavioral effects. For example, a seemingly "simple" vitamin B12 deficiency may produce a wide array of symptoms, including memory deficits, mood disturbances, anemia, numbness, and even psychosis.

Consequently, the persistent search for a universal and multifaceted explanation of stuttering—motivated by the amount of biological and cognitive differences among PWS—has led researchers to conceptualize the nature and causes of stuttering in increasingly complicated ways. While such models aim to capture the disorder's full richness, they have often proven difficult to test or validate scientifically, ultimately hindering our understanding of stuttering.

By interpreting the wide range of biological and cognitive differences observed in stuttering as indicative of cognitive heterogeneity, we have adopted a perspective that enables a more focused and tractable approach. This perspective enables us to define a logical and causal framework (see Section 1 of Chapter 1) that restricts investigation to the cognitive level, emphasizing the search for direct or mechanical (proximal) causes that are necessary and/or sufficient for stuttering to occur—those factors immediately responsible for stuttering-like behaviors. This perspective has also enabled us to differentiate between factors that could

contribute to stuttering and those that are less likely (see Chapter 4). For example, our approach suggests that the role of tactile perception may have been previously underestimated in the literature—a point we will explore in greater detail in Section 4.

Although exploring the biological level and identifying indirect (distal) etiologies would eventually provide a more comprehensive account of stuttering, we contend that it is not a prerequisite for acquiring useful causal knowledge or for developing effective, targeted interventions.

By examining stuttering from a different perspective, we believe we have added important background knowledge that yields several key implications and serves as a prerequisite for meaningful progress in the search of causal knowledge (Danks & Davis, 2020).

2.2 The causal implication of cognitive heterogeneity

In Chapter 4, we discussed how the direct demonstration of cognitive heterogeneity among PWS carries significant theoretical, clinical, and methodological implications. The aim here is not to reiterate those points, but rather to examine how the recognition of cognitive heterogeneity must inform and guide the pursuit of causal knowledge about stuttering.

In the context of the human mind, searching for causal relationships is particularly challenging because the underlying mechanisms are often neither directly observable nor fully understood. Moreover, these mechanisms are not always easily testable. In the context of a developmental disorder that is characterized by cognitive heterogeneity, determining a causal relationship becomes even more challenging. Disentangling what constitutes a cause, a consequence, a compensatory mechanism, or an unrelated comorbid condition is a daunting task. This complexity raises an important question: if uncovering causal relationships seems so daunting, how should we proceed? We propose that multi-faceted methodological approaches are necessary.

First, at the individual level and in parallel with the approach adopted here, CASE STUDIES can provide valuable insight. By administering a comprehensive battery assessing the range of cognitive abilities required for fluent speech, it is possible to generate a detailed

cognitive profile for an individual who stutters. This enables clinicians and researchers to identify specific areas of weakness and strengths, and to interpret these in light of a theoretical model of fluent speech. Based on this profile, targeted (longitudinal) interventions can then be designed and implemented. If such interventions prove effective in reducing stuttering-like behaviors, this will support the hypothesized locus of the disorder for that particular individual. Reporting well-documented case studies of this kind would not only benefit clinical practice, but would also be informative for the broader scientific community, contributing incrementally to our overall causal understanding of stuttering. This perspective on causation is supported by the philosophical theory of interventions (see Woodward, 2003). As Woodward (2003, p. 10) states, "We have at least the beginning of an explanation when we have identified factors or conditions such that manipulations or changes in those factors or conditions will produce changes in the outcome being explained". However, there are important limitations to causal inferences that can be drawn from intervention studies. A key assumption underlying such interventions is that, when manipulating a factor presumed to be causally related to an outcome, all other relevant factors must be held constant, meaning the intervention is expected to influence only its intended target (Eberhardt & Scheines, 2016; Woodward, 2003). In practice, even the most carefully designed interventions cannot completely "freeze" other cognitive abilities and mental state (Eronen, 2020). This makes it difficult to determine with certainty whether the intervention truly targeted only the specific ability under investigation. Such challenges are especially pronounced in cases that involve a wide array of cognitive difficulties, where overlapping and interacting domains of cognition increase the risk of unintended effects beyond the focus of the intervention.

Second, at the group level, MACHINE LEARNING TECHNIQUES have become increasingly prominent in recent years. These approaches encompass a broad set of methods designed to uncover patterns and associations within independent and identically distributed data sets (see Kaddour et al., 2022). Among these, clustering methods are particularly valuable for their ability to partition data into distinct groups, ensuring that observations within each group are similar to one another, while those in different groups are substantially different (James et al., 2021). Clustering methods offer important advantages for investigating cognitive heterogeneity (see Carruthers et al., 2019; Gilbert et al., 2013; Giofrè et al., 2019). By identifying data-driven subgroups within a population, these approaches enable researchers to group individuals who share similar cognitive strengths and

weaknesses. This moves research beyond the presumption of a single, homogeneous disorder, instead revealing distinct cognitive profiles across individuals. Such subgroup identification can better inform the development of targeted assessments and interventions, and can ultimately facilitate the differentiation between primary causes, compensatory mechanisms, and comorbid conditions within a heterogeneous population. However, clustering analyses also have significant limitations (see James et al., 2021). Among these is that researchers must make several methodological decisions prior to analysis, each of which can substantially influence the resulting groups. For example, in K-means clustering, the number of clusters must be specified in advance, often requiring a degree of background knowledge or theoretical justification about the underlying structure of the data. Moreover, clustering methods generally require large amounts of high-quality data, and the outcomes are highly dependent on the variables selected for analysis. Different clustering algorithms and parameter choices may yield divergent subgroupings, necessitating the comparison of several analytic methods to identify the most meaningful or interpretable structure. Even then, individuals within the same cluster may still display considerable variation. In summary, while clustering is a powerful tool for exploring and characterizing heterogeneity within cognitive and clinical data, its effectiveness is contingent on thoughtful methodological decisions and a strong understanding of the context being studied. Rather than yielding definitive answers, clustering exposes potentially interesting patterns for further investigation, it requires the researcher to consider any solution as it exposes some interesting aspects of the data.

Third, LONGITUDINAL DESIGNS represent powerful, mixed individual- and group-level methodological approaches for advancing causal knowledge in the study of developmental disorders (see Karmiloff-Smith, 2012). By assessing biological, cognitive, and behavioral changes from infancy onward, longitudinal studies enable researchers to trace developmental trajectories over time. This design is particularly effective for mapping the early presentation of symptoms to their potential underlying mechanisms and for observing how these factors evolve. Through the examination of cohorts, it becomes possible to identify patterns and associations of symptoms shared within a population, while also, through detailed case analyses, elucidating the unique developmental paths of individuals—facilitating the identification of cognitive subtypes. Furthermore, by observing the temporal sequence of changes, longitudinal research allows for the determination of the order in which symptoms

emerge and develop, thus helping to disentangle direct causes from consequences, compensatory mechanisms, or comorbidities. By enabling a nuanced exploration of the timing, interaction, and interdependence of symptoms and mechanisms at both the group and individual levels, longitudinal studies offer a particularly effective approach for unraveling the complexity of developmental disorders. Interestingly, this approach aligns with the perspective proposed by Attanasio and collaborators (1998), who suggested that understanding stuttering requires investigating the proximal causes of stuttering, specifically, the repetition of speech elements, typically occurring at the beginning of an utterance. While longitudinal approaches provide valuable insights into developmental disorders, they are accompanied by considerable challenges. These studies are highly time-consuming and often place significant demands on both parents and children, sometimes resulting in unnecessary stress for participating families (Annaz, Karmiloff-Smith, & Thomas, 2008). In addition, it is inherently difficult to assess children directly at the very onset of their disorder, as early manifestations may be subtle or overlooked, and the timing of initial assessment often depends on when symptoms become sufficiently pronounced to attract attention.

Fourth, at the technical level, COMPUTATIONAL MODELING offers a powerful means for causal inferences in the study of cognitive abilities and developmental disorders. Mathematical and algorithmic models can simulate mental processes to explain, predict, and help understand how specific cognitive functions operate in humans (Farrell & Lewandowsky, 2018). For example, the GODIVA model, presented in Chapter 1 (Section 2.2), is a neurocomputational model that captures speech motor production and control, simulating the transformation of speech sounds into articulatory movements (Guenther, 1994; Guenther, 2016; Meier & Guenther, 2023; Tourville & Guenther, 2011). This model has provided mechanistic insights into disorders such as stuttering, demonstrating that specific impairments within its architecture can reproduce the characteristic symptoms of the disorder (see Chang & Guenther, 2020; Civier et al., 2010, 2013; Guenther, 2016). When a model can show that introducing targeted deficits leads to the patterns of behavior or symptoms seen in affected individuals, it constitutes strong evidence for the causal relevance of the simulated mechanism. Connectionist computational models, in particular, are valuable for disentangling the sources of cognitive heterogeneity in developmental disorders. As discussed by Thomas (2003) and Thomas & Karmiloff-Smith (2002), these models allow researchers to explore whether observed behavioral diversity within a disorder group arises

from a single underlying cognitive cause—with variability due to individual developmental differences, compensation, or noise—or whether the heterogeneity reflects multiple distinct underlying causes, each generating its own behavioral profile. By systematically manipulating parameters and constraints within the model, researchers can simulate both situations. In a homogeneous scenario, where all models share the same constraint, behavioral variability is generally limited and tends to decrease over time, as the group's patterns converge. In contrast, when different constraints are introduced across models, as would be the case for a heterogeneous group, the resulting behavioral profiles remain diverse, and variability across tasks persists or increases, reflecting the influence of multiple causal pathways. These simulations are especially informative when compared to empirical data from actual clinical populations. If real-world behavioral profiles mirror those from homogeneous models, a single-cause theory gains support. Conversely, persistent or increasing heterogeneity across multiple tasks, as seen in heterogeneous model simulations, suggests that the disorder may in fact arise from multiple independent mechanisms. Importantly, this approach can also give insight into how development and intervention might affect different subgroups, providing a rigorous theoretical basis for designing tailored interventions that match the precise source of an individual's difficulties. However, these methods are deeply dependent on our current knowledge of the cognitive and neural organization of the human mind and can be difficult to implement. Their effectiveness relies on the quality of theoretical and background knowledge available regarding development and developmental disorders. As a result, computational models face ongoing challenges in accurately linking model architecture to real-world phenomena, which can lead to misinterpretation or oversimplification of results.

Demonstrating that AWS display distinct cognitive profiles across the abilities that support fluent speech is a crucial step toward a better understanding of the physiopathology of stuttering. By establishing that AWS exhibit diverse profiles of cognitive strengths and weaknesses, we can begin to identify factors that may be associated to stuttering and distinguish those less likely to have a causal role. This, in turn, enables the refinement of theoretical, clinical, and methodological assumptions, and guides the adoption of best practices for making causal inferences, allowing us to more effectively disentangle causes from compensatory mechanisms, consequences, and comorbidities.

Surprisingly, our findings also reveal a homogeneous difficulty across all AWS tested: a consistent orofacial tactile deficit. This deficit remained stable even upon retesting. In the following section, we will examine how this shared orosensory deficit might be related to stuttering, either as a possible causal mechanism or, alternatively, as a consequence of the disorder.

3. Orosensory tactile weaknesses and stuttering: a cause or a consequence?

The second main finding of the present thesis was the unexpected discovery that all participants exhibited an orofacial tactile perception deficit (See Chapter 4 and Appendix discussion of the consistency of these difficulties). Notably, in some individuals, this orosensory weakness was the only impairment detected, suggesting it may represent both a necessary and sufficient condition for the development of stuttering. As discussed in Chapter 1 (see Sections 2.2 and 3.1.3), sensory information—both auditory and somatosensory—plays a fundamental role in the acquisition, maturation, and ongoing updating of speech motor control (Bradshaw et al., 2021; Guenther, 2016). In Chapter 4, we proposed three principal hypotheses to explain how difficulties in detecting or discriminating tactile stimuli could causally contribute to stuttering-like behaviors. In this section, we will first present additional theoretical arguments that further strengthen the potential causal link, underlining the need to pay greater attention to the role of tactile impairments in the emergence and persistence of stuttering. We then discuss the alternative perspective that such impairments could also be a consequence of stuttering.

3.1 Orosensory tactile weaknesses as a potential cause of stuttering

In Chapter 4, we discussed three hypothetical perspectives on how a subtle orofacial tactile deficit might contribute to stuttering within the GODIVA framework. Here, we aim to address how the specific tactile difficulties identified in this thesis could (1) contribute to the emergence of stuttering, as conceptualized by Max et al. (2004), and (2) disrupt the timing network that supports fluent speech production—thereby increasing the risk of stuttering-like behaviors.

3.1.1 The potential role of tactile difficulties in the emergence of stuttering

Before addressing the two perspectives outlined by Max et al. (2004) on how stuttering might emerge, I will first present key theoretical arguments concerning speech motor learning. Although these foundational points have already been discussed in Chapter 1 (see Section 2.2.1), they will be elaborated upon here in greater detail at each stage. This expanded approach is intended to facilitate clearer and more direct parallels with the theoretical account of stuttering advanced by these authors.

Emergence of stuttering due to unstable or insufficiently activated internal models.

Acquiring fluent speech requires infants to map the sensory representations (auditory and somatosensory) of each speech unit of their native language to the motor representations needed to produce those sounds (see Meier & Guenther, 2023). During the babbling phase, infants generate a variety of speech-like movements, receiving real-time feedback from the sounds they hear and the sensations they feel as a result. Although these early productions may appear random, they are essential for establishing the foundational links between articulatory actions and their sensory outcomes.

As development progresses, infants increasingly attempt to imitate the speech sounds of their caregivers. In doing so, they work to match the sensory outcomes they observe—how a speech sound should feel and sound—with the articulatory movements necessary to reproduce them. This mapping process relies on constructing and refining what are known as inverse models: neural mechanisms that translate a desired sensory target into the corresponding motor commands (Bradshaw et al., 2021; Meier & Guenther, 2023; Guenther 2016). In parallel, infants continually compare the actual sensory feedback from their own productions to their intended targets. Discrepancies between expected and actual outcomes—so-called sensory errors—are detected using feedback control mechanisms. These errors drive real-time corrections and, crucially, are used to update and improve the internal (feedforward) motor commands responsible for generating future attempts. Because early attempts at speech production are often imprecise, infants rely heavily on sensory feedback to guide and correct their movements. Over time and with repeated practice, successful productions become more frequent, and the associated sensorimotor representations are

strengthened. As these internal models (feedforward) become more accurate, infants can anticipate the necessary movements to produce a given speech sound without needing to rely on feedback for each attempt. This developmental process gradually shifts speech control from a predominantly feedback-dependent mode to a more efficient and rapid feedforward control system (Bradshaw et al., 2021; Meier & Guenther, 2023; Guenther 2016).

If a child has subtle sensory difficulties—such as deficits in orofacial tactile perception—the acquisition, retention and updating of the internal mappings between sensory targets and articulatory movements may become weak or inconsistent. According to Max and collaborators (2004), when inverse models that translate sensory goals into motor commands are inadequately developed, the resulting feedforward commands may be poor or faulty. When these inaccurate motor commands are executed, the resulting sensory feedback will likely diverge from the intended target. As a consequence, frequent feedback-based corrections become necessary, often leading to interruptions or resets of the feedforward command sequence (Bradshaw et al., 2021). This process can manifest behaviorally as repetitions or prolongations of sounds and syllables. As Bradshaw et al. (2021) describe Max et al. (2004)'s theory, they stated: "in this account, stuttering is proposed to involve an impairment in the ability to use auditory feedback early in development to establish [such] bidirectional mappings between motor commands and sensory consequences." While their emphasis is on auditory feedback, similar impairments in somatosensory (including tactile) feedback could plausibly have equivalent disruptive effects on the development of these sensorimotor mappings, potentially contributing to the onset and persistence of stuttering.

It is interesting to note that Guenther (2016, p.115) proposes an estimated developmental timeline for the main components of the speech motor system. Specifically, somatosensory target maps—which specify the expected somatosensory outcomes of speech sounds and are derived from the internal model (see Chapter 1, Section 2.2.2)—undergo significant maturation between approximately 1.5 and 3.5 years of age. This period closely aligns with the typical window for the onset of developmental stuttering. Such a parallel suggests that incomplete or unstable development of these somatosensory target maps may result in weak or unreliable feedforward motor commands, contributing to the emergence of stuttering during this critical phase of speech motor development.

Emergence of stuttering due to overreliance on feedback control.

Max and collaborators (2004) offered an additional perspective on the development of stuttering. As discussed previously, infants initially rely heavily on sensory feedback to acquire, control, correct, retain, and update their sensorimotor commands for speech. With time and repeated practice, successful productions become more frequent, gradually strengthening the underlying sensorimotor representations. As these internal models increase in accuracy, infants become increasingly able to anticipate and execute the movements required for specific speech sounds without depending on feedback for each attempt. This developmental trajectory reflects a shift from predominantly feedback-dependent control to a more efficient and rapid feedforward control system (Bradshaw et al., 2021; Meier & Guenther, 2023; Guenther, 2016). This means that early in development, sensory targets for speech sounds are broad, and children's systems recognize a wide range of productions as "correct"—that is, they maintain a high threshold for triggering sensory error-based motor resets. This inherent flexibility is adaptive: it protects developing speakers from halting or repeating themselves over minor errors as they experiment with new sounds. Over time, as their skills and feedforward control improve, these sensory targets become more precise, and the system grows increasingly sensitive to small discrepancies between intended and actual speech (Max et al., 2004).

Max et al. (2004) suggest that stuttering stems from an inability, early in development, to achieve the typical transition from feedback control to feedforward control. As a result, these children remain overreliant on feedback systems to guide their speech. This overdependence on feedback introduces instability in the speech motor control system, because there is an inherent time lag between issuing a motor command and receiving its associated sensory (auditory and somatosensory) feedback (Bradshaw et al., 2021).

According to Max et al. (2004), in children who stutter, the feedforward commands that should gradually become more accurate remain weak or underdeveloped. Consequently, the magnitude of sensory error signals—differences between intended and actual speech output—does not decrease as it should over time. Simultaneously, as the speech system matures, its sensitivity to error naturally increases: the threshold for triggering a motor reset (an interruption or correction) becomes progressively lower. For children who continue to rely on feedback, persistent large sensory errors increasingly exceed this lowered threshold,

resulting in frequent error detection and triggering repeated motor corrections or resets. This mechanism gives rise to the characteristic repetitions and prolongations observed in stuttering. This process is further supported by computational simulations. Civier et al. (2010) demonstrated that overreliance on feedback leads to the accumulation of large sensorimotor errors, mainly because the time delays inherent in feedback processing hinder timely error correction. These errors are always present to some extent, but when they surpass the critical threshold monitored by the system, corrective actions—such as interruptions or restarts—are initiated.

Considering the tactile difficulties observed in this thesis, we can hypothesize that tactile sensory information may be less reliable or require longer processing times, as the system needs more time to acquire sufficient tactile input. As a result, the child's speech motor system may take longer to accumulate adequate somatosensory evidence for evaluating the accuracy of their speech output. This extended processing period further prolongs the feedback loop, thus exacerbating the inherent delay between the generation of a motor command and the perception of its sensory consequences. Consequently, corrections and adjustments become less timely and potentially more disruptive, undermining speech fluency. Alternatively, if tactile feedback is insufficient or unreliable, the system may compensate by increasing its reliance on auditory feedback. However, because auditory feedback is naturally subject to even longer processing delays, this compensation can further slow down the feedback control loop. Both scenarios—prolonged somatosensory processing and increased dependence on slower auditory feedback—accentuate the temporal mismatch that is a core feature of feedback-biased control models. This amplified delay increases the likelihood of speech interruptions, resets, or repetitions, thereby elevating the risk of disfluent speech behaviors.

3.1.2 The potential role of tactile difficulties in the disruption of the timing network that supports fluent speech

We have discussed in Chapter 1 (Section 3.1.4) that sensory information—specifically auditory and somatosensory inputs—plays a crucial role in the brain's "timing network" (Chang & Guenther, 2020). By continuously monitoring and integrating these sensory modalities, the *basal ganglia–premotor cortex loop* ensures that selected motor commands for speech are both initiated and terminated precisely at the right moments, thereby

supporting the smooth and sequential production of speech units (Bohland et al., 2010). Within this framework, stuttering has been hypothesized to arise from abnormal sensorimotor integration, wherein sensory information is not combined efficiently—such as when auditory and somatosensory signals are processed at different speeds, or when the sensory context fails to match the state required to launch the next motor program (Bradshaw et al., 2021). These temporal mismatches disrupt the internal timing network's ability to accurately coordinate the onset and offset of speech motor sequences and to effectively advance through the speech sequence, ultimately hindering fluent speech production (Sares et al., 2019; Slis et al., 2023). Civier et al. (2013) demonstrated through computational simulations that disruption of the *basal ganglia–premotor cortex loop* delays the initiation of the next syllable's motor program, resulting in stuttering-like speech behaviors.

Considering the tactile difficulties observed in this thesis, it is possible that imprecise tactile perception hinders the *basal ganglia–premotor cortex loop* from accurately determining the current position of the articulators, due to reduced somatosensory input. As a result, the system may lack precise information about articulator placement, which is essential for the smooth sequencing and initiation of motor commands during speech. Alternatively, reduced perception of contact pressure between articulators may necessitate a longer period to acquire sufficient tactile information. This delay could introduce a timing gap between the arrival of auditory and somatosensory feedback, thereby complicating multisensory-motor integration (see Slis et al., 2023). Such temporal mismatches may dysregulate the internal timing network, compromising its ability to coordinate the timely initiation and termination of speech motor programs, and ultimately leading to disruptions in speech fluency.

To summarize, an orofacial tactile deficit, such as a subtly elevated tactile threshold, could contribute to the emergence of stuttering-like behaviors through two primary mechanisms: (1) By weakening sensorimotor integration early in development, such a deficit would lead to the formation of weak and inaccurate feedforward motor commands. An inability to effectively rely on feedforward control necessitates overreliance on feedback-based control, which is inherently slower and less stable, thereby increasing the risk of disruptions in speech fluency. (2) A tactile deficit could also impair the neural timing network responsible for the smooth and sequential production of speech. Disruption of this timing

system compromises the precise initiation and termination of motor programs, further hindering fluent speech production.

3.2 Orosensory tactile weaknesses as a potential consequence of stuttering

So far, different perspectives have been discussed to explain how subtle weaknesses in orosensory processing could potentially contribute to the emergence of stuttering. However, it is equally plausible that these difficulties are not a root cause of stuttering but instead develop as a consequence of living with the disorder. In the next section, we will discuss two possible explanations for how orosensory weaknesses may arise as a result of stuttering itself.

Muscle tension is frequently described as a symptom of stuttering. Alm and collaborators (2025) propose that excessive tension may arise because PWS engage their muscles more forcefully in an effort to break through a fixed articulatory posture and resume speech. This excessive and chronic muscle tension can influence sensory perception. For example, Cohen and Starr (1985) have shown that during active muscle contraction, there is a central gating or filtering of proprioceptive and tactile inputs; this process helps prevent the brain from becoming overloaded with sensory signals when muscles are actively engaged. If PWS frequently experience muscle tension in the orofacial region, this sustained tension could plausibly lead to a chronic reduction in orosensory perception. Over time, such central sensory inhibition may make it more difficult for PWS to detect and interpret tactile signals from their oral structures, contributing to a persistent orofacial tactile deficit.

Another possibility is that PWS may practice speech less frequently to avoid stuttering episodes, as behavioral avoidance and reduced communication are commonly reported among PWS (Bauerly, 2024). However, practice is foundational for acquiring and refining somatosensory representations, as these representations can only be developed and maintained through active speech experience (Meier & Guenther, 2023). By producing less speech than non-stutterers, it is possible that the tactile perception of PWS is less precise, which may further affect their speech motor control. For example, studies in other motor domains have shown that force control, such as thumb pressure, improves with practice in learning tasks; participants who become more efficient at applying the optimal amount of

force also achieve higher performance. This highlights not only the importance of practice for acquiring precise tactile representations, but also underscores how refined tactile information actively supports motor learning and execution (Cienfuegos et al., 2024).

The paradox observed in this thesis points to several unanswered questions for future research. Are orosensory difficulties the core mechanism that triggers stuttering, or are they consequences of the disorder's impact over time? Do the variable cognitive weaknesses seen among AWS reflect compensatory adaptations, unrelated individual characteristics, co-occurring conditions, or other individual factors that are superimposed on the core disorder, thus accounting for the behavioral variability observed across PWS? Addressing these questions will require rigorous experimental and longitudinal studies, that will pave the way for theory-driven interventions and more individualized clinical care.

4. Limits and future perspectives

Several lines of inquiry for clinical practice and research can be identified based on the findings of the present thesis.

First, current clinical evaluations of stuttering lack appropriate tools for assessing the potential cognitive factors that may underlie the disorder. Typically, stuttering assessments focus primarily on identifying the presence of stuttering and categorizing its severity. While accurate diagnosis is essential, such approaches offer limited guidance for clinicians in tailoring therapeutic interventions to individual needs. Consequently, it is not surprising that many clinicians report feeling unprepared or uncomfortable when it comes to treating stuttering (see Yaruss, 1997), in part because they lack proper assessment tools like those available for other developmental disorders such as dyslexia or dyscalculia. This need is amplified by the cognitive heterogeneity observed among PWS, underscoring the importance of enabling clinicians to assess the specific cognitive, motor and sensory profiles of each individual patient.

In this thesis, we have taken an initial step toward addressing this gap by developing the first cognitive assessment battery designed specifically for stuttering. Although this battery is not yet ready for clinical implementation and will require further refinement, it provides a foundation for future advances. For example, improvements are needed to enhance

the assessment of somatosensory abilities, particularly in optimizing task design to balance precision with efficiency. In addition, technical modifications are necessary to allow reliable evaluation of multisensory-motor integration.

Second, it is also important to note the scope and limitations of the battery design in the present thesis. The assessment tools developed here were guided by specific theoretical hypotheses aimed at explaining stuttering. Two critical considerations arise from this approach. First, the battery was not designed to exhaustively assess the all spectrum of abilities of one specific domain that is potentially linked to stuttering; therefore, the absence of a difficulty in one specific assessment does not conclusively rule out that domain as a causal contributor for a given individual—rather, it suggests that the particular aspect of that domain tested is unlikely to contribute to that patient's stuttering. Second, the creation of cognitive and sensorimotor assessments is intrinsically constrained by our current theoretical understanding of stuttering (and speech production). This creates something of a vicious circle: without solid theories, it is challenging to design truly appropriate assessment tools, yet it is through observing divergences and patterns in cognitive and sensorimotor assessment that we are able to refine and build those very theories. Ultimately, progress depends on the iterative interplay between empirical observation and theoretical development.

Third, large-scale studies are essential to obtain broader and more representative datasets, which are crucial for identifying potential subtypes of stuttering. However, conducting in-depth cognitive assessments on a large number of individuals is inherently time-consuming and logistically demanding, particularly at the scale of a thesis, posing significant challenges for researchers in this field. During the course of this thesis, we encountered substantial difficulties in recruiting a sufficient number of participants, reflecting a common barrier in any study requiring the assessment of patient populations.

To overcome these obstacles and enable more meaningful scientific progress, there is a strong need for coordinated international collaboration focused on uncovering the causal mechanisms of stuttering. Establishing an international research consortium, analogous to the collaborative spirit seen in the field of particle physics, such as CERN (Zalewska, 2014), could facilitate the pooling of resources, data, and expertise across institutions and countries. Such an initiative would not only accelerate participant recruitment and data collection but

would also enhance the diversity and generalizability of findings, ultimately advancing our understanding of stuttering subtypes (if they exist) and their underlying mechanisms.

Fourth**,** the focus of this thesis has been on the cognitive level. To fully understand the mechanisms underlying stuttering, future research must draw explicit links between cognitive processes and their biological substrates. However, we contend that meaningful and accurate mapping between these two levels first depends on a much deeper understanding of the behavioral and cognitive architecture of stuttering, including the identification of potential subtypes and the disentanglement of causes, consequences, compensation strategies, and comorbid factors. There is a temptation in modern neuroscience to seek the neural circuits and brain mechanisms of cognitive function and disorders directly (Krakauer et al., 2017). Yet, as argued by Krakauer and colleagues (2017), understanding brain function and dysfunction requires detailed behavioral and cognitive analysis as both a prerequisite and a complement to neuroscientific investigation. Without such background, biological findings are at risk of being misinterpreted.

To illustrate this point, consider a classic analogy from computer science: if the brain is the hardware (processors), and cognitive processes are the software (programs or algorithms), can we discover what the brain is actually doing simply by studying its physical structure (the hardware)? Or must we first understand the "software", the psychological processes and computational rules that govern behavior (Krakauer et al., 2017)? This analogy has been directly tested by Jonas and Kording (2017), who applied standard neuroscience data analysis methods to a well-understood microprocessor. By exhaustively mapping the chip's connections, analyzing the effects of removing individual transistors, and recording local activity patterns, all with access to "unlimited" data, they found that these approaches were nonetheless insufficient for truly understanding the functional computations of the processor; specifically, they could not explain how the microprocessor executed or performed its programmed tasks (such as running video games). This work underscores the fundamental limitation of studying hardware in isolation from software: without a rigorous understanding of the cognitive and behavioral "programs" being implemented, even the most comprehensive biological analyses can yield only superficial or incomplete explanations.

5. Conclusion

Through a comprehensive assessment of the cognitive abilities underlying fluent speech, this thesis aimed to clarify the patterns of strengths and weaknesses experienced by PWS within these domains. Using a multiple single-case approach, two complementary studies provided in-depth examinations of the cognitive abilities of AWS. This work revealed the following pattern: individual cognitive heterogeneity coexists with shared weaknesses in orosensory tactile perception.

The shared orosensory tactile weaknesses observed in all participants points to a potentially crucial—yet previously underestimated—role for tactile processing in stuttering. Whether this deficit is a root cause or a secondary consequence of the disorder remains unresolved; this opens important avenues for future research.

Moreover, the cognitive diversity identified in AWS highlights the limitations of group-level analyses and reinforces the need for comprehensive individual assessments in both research and clinical contexts to better capture the diversity of mechanisms underlying stuttering. Future studies and interventions will benefit from integrating detailed cognitive and sensorimotor profiling to enable a more nuanced understanding of individual differences, etiological mechanisms, and tailored treatment strategies.

In sum, these findings advocate for a more comprehensive evaluation of the components linked to fluent speech in PWS, both in research and clinical settings. Adopting such an approach will improve both scientific and clinical understanding of stuttering's diversity and guide the development of more individualized, theory-driven interventions.

References

Abbs, J., & Gracco, V. (1983). Sensorimotor actions in the control of multi-movement speech gestures. *Trends in Neurosciences, 6*(C), 391–395. https://doi.org/10.1016/0166-2236(83)90173-X

Abbs, J., & Gracco, V. (1984). Control of complex motor gestures: Orofacial muscle responses to load perturbations of lip during speech. *Journal of Neurophysiology, 51*(4), 705–723. https://doi.org/10.1152/jn.1984.51.4.705

Adams, M., & Ramig, P. (1980). Vocal characteristics of normal speakers and stutterers during choral reading. *Journal of Speech and Hearing Research, 23*(2), 457–469. https://doi.org/10.1044/jshr.2302.457

Alderson-Day, B., & Fernyhough, C. (2015). Inner speech: Development, cognitive functions, phenomenology, and neurobiology. *Psychological Bulletin, 141*(5), 931–965. https://doi.org/10.1037/bul0000021

Alfonso, P. (1991). Subject definition and selection criteria for stuttering research in adult subjects. *Haskins Laboratories Status Report on Speech Research, 105/106*, 231–242.

Allen, G. (1973). Segmental timing control in speech production. *Journal of Phonetics, 1*(3), 219–237. https://doi.org/10.1016/S0095-4470(19)31387-7

Alm, P. (2004). Stuttering and the basal ganglia circuits: A critical review of possible relations. *Journal of Communication Disorders, 37*(4), 325–369. https://doi.org/10.1016/j.jcomdis.2004.03.001

Alm, P. (2005). *On the causal mechanisms of stuttering* [Doctoral dissertation, Lund University]. Lund University Publications.

Alm, P., Brösemyr, T., Grinde, S., Johansson, S., Karlsson, J., Nordgren, G., ... & White, D. (2025). Stuttering: the nature of the speech disruptions—a multimodal study of articulation and phonation. *Frontiers in Human Neuroscience*, *19*, 1623308. https://doi.org/10.3389/fnhum.2025.1623308

Ambrose, N. (2004). Theoretical perspectives on the cause of stuttering. *Contemporary Issues in Communication Science and Disorders, 31*, 80-91. https://doi.org/10.1044/cicsd_31_S_80

Ambrose, N., Yairi, E., Loucks, T. M., Seery, C., & Throneburg, R. (2015). Relation of motor, linguistic and temperament factors in epidemiologic subtypes of persistent and recovered stuttering: Initial findings. *Journal of Fluency Disorders, 45*, 12–26. https://doi.org/10.1016/j.jfludis.2015.05.004

American Speech-Language-Hearing Association. (2005). (Central) auditory processing disorders—The role of the audiologist. https://www.asha.org/policy/ps2005-00114/

Anderson, J. (2007). Phonological neighborhood and word frequency effects in the stuttered disfluencies of children who stutter. *Journal of Speech, Language, and Hearing Research, 50*(1), 229–247. https://doi.org/10.1044/1092-4388(2007/018)

Anderson, J., & Conture, E. (2004). Sentence-structure priming in young children who do and do not stutter. *Journal of Speech, Language, and Hearing Research, 47*(3), 552–571. https://doi.org/10.1044/1092-4388(2004/043)

Anderson, J., & Ofoe, L. (2019). The role of executive function in developmental stuttering. *Seminars in Speech and Language, 40*(4), 305–319. DOI: 10.1055/s-0039-1692965

Anderson, J., & Wagovich, S. (2010). Relationships among linguistic processing speed, phonological working memory, and attention in children who stutter. *Journal of Fluency Disorders, 35*(3), 216–234. https://doi.org/10.1016/j.jfludis.2010.04.003

Anderson, J., Wagovich, S., & Hall, N. (2006). Nonword repetition skills in young children who do and do not stutter. *Journal of Fluency Disorders, 31*(3), 177–199. https://doi.org/10.1016/j.jfludis.2006.05.001

Andrews, G., Craig, A., Feyer, A., Hoddinott, S., Howie, P., & Neilson, M. (1983). Stuttering: A review of research findings and theories circa 1982. *Journal of Speech and Hearing Disorders, 48*(3), 226–246. https://doi.org/10.1044/jshd.4803.226

Andrews, G., Howie, P., Dozsa, M., & Guitar, B. (1982). Stuttering: Speech pattern characteristics under fluency-inducing conditions. *Journal of Speech and Hearing Research, 25*(2), 208–216. https://doi.org/10.1044/jshr.2502.208

Angelidis, A., Solis, E., Lautenbach, F., van der Does, W., & Putman, P. (2019). I'm going to fail! Acute cognitive performance anxiety increases threat-interference and impairs WM performance. *PLOS ONE, 14*(2), e0210824. https://doi.org/10.1371/journal.pone.0210824

Annaz, D., Karmiloff-Smith, A., & Thomas, M. S. (2008). The importance of tracing developmental trajectories for clinical child neuropsychology. *Child Neuropsychology, 14*(6), 495–518.

Archibald, L., & De Nil, L. (1999). The relationship between stuttering severity and kinesthetic acuity for jaw movements in adults who stutter. *Journal of Fluency Disorders, 24*(1), 25–42. https://doi.org/10.1016/S0094-730X(98)00023-0

Armson, J., Kiefte, M., Mason, J., & Croos, D. (2006). The effect of SpeechEasy on stuttering frequency in laboratory conditions. *Journal of Fluency Disorders, 31*(2), 137–152. https://doi.org/10.1016/j.jfludis.2006.04.004

Arndt, J., & Healey, E. (2001). Concomitant disorders in school-age children who stutter. *Language, Speech, and Hearing Services in Schools, 32*(2), 68–78. https://doi.org/10.1044/0161-1461(2001/006)

Asal, S., & Abdou, R. (2014). The study of central auditory processing in stuttering children. *The Egyptian Journal of Otolaryngology, 30*(4), 357–362.

Aschersleben, G. (2002). Temporal control of movements in sensorimotor synchronization. *Brain and Cognition, 48*(1), 66–79. https://doi.org/10.1006/brcg.2001.1304

Assaneo, M., Ripollés, P., Tichenor, S., Yaruss, J., & Jackson, E. (2022). The relationship between auditory-motor integration, interoceptive awareness, and self-reported stuttering severity. *Frontiers in Integrative Neuroscience, 16*, Article 869571. https://doi.org/10.3389/fnint.2022.869571

Association Parole Bégaiement. (n.d.). Le bégaiement. https://www.begaiement.org/begaiement/

Attanasio, J., Onslow, M., & Packman, A. (1998). Representativeness reasoning and the search for the origins of stuttering: A return to basic observations. *Journal of Fluency Disorders, 23*(4), 265–277. https://doi.org/10.1016/S0094-730X(98)00012-6

Bachy, S., Dister, A., Francard, M., Geron, G., Giroul, V., Hambye, P., Simon, A.-C., & Wilmet, R. (2004). Conventions de transcription régissant les corpus de la banque de données VALIBEL. Université Catholique de Louvain.

Baddeley, A. (2003). Working memory: Looking back and looking forward. *Nature Reviews Neuroscience, 4*(10), 829–839. https://doi.org/10.1038/nrn1201

Bakhtiar, M., Abad, A., & Sadegh, S. (2007). Nonword repetition ability of children who do and do not stutter and covert repair hypothesis. *Folia Phoniatrica et Logopaedica, 59*(3), 160–167.

Barbarotto, R., Laiacona, M., Macchi, V., & Capitani, E. (2002). Picture reality decision, semantic categories and gender: A new set of pictures, with norms and an experimental study. *Neuropsychologia, 40*(10), 1637–1653. https://doi.org/10.1016/S0028-3932(02)00029-5

Barbosa, P. (2007). From syntax to acoustic duration: A dynamical model of speech rhythm production. *Speech Communication, 49*(9), 725–742. https://doi.org/10.1016/j.specom.2007.04.013

Baxter, S., Johnson, M., Blank, L., Cantrell, A., Brumfitt, S., Enderby, P., & Goyder, E. (2015). The state of the art in non-pharmacological interventions for developmental stuttering. Part 1: A systematic review of effectiveness. *International Journal of Language & Communication Disorders, 50*(5), 676–718. https://doi.org/10.1111/1460-6984.12171

Becker, A. (2023). The benefits of single-subject research designs and multi-methodological approaches for neuroscience research. *Frontiers in Human Neuroscience, 17*, Article 1190412. https://doi.org/10.3389/fnhum.2023.1190412

Belyk, M., Kraft, S., & Brown, S. (2015). Stuttering as a trait or state – an ALE meta-analysis of neuroimaging studies. *European Journal of Neuroscience, 41*(2), 275–284. https://doi.org/10.1111/ejn.12765

Bernard, R., Hofslundsengen, H., & Frazier Norbury, C. (2022). Anxiety and depression symptoms in children and adolescents who stutter: A systematic review and meta-analysis. *Journal of Speech, Language, and Hearing Research, 65*(2), 624–644. https://doi.org/10.1044/2021_JSLHR-21-00236

Berry, J. (2011). Speaking rate effects on normal aspects of articulation: Outcomes and issues. *Perspectives on Speech Science and Orofacial Disorders, 21*(1), 15–26. https://doi.org/10.1044/ssod21.1.15

Bishop, D. (1997). Cognitive neuropsychology and developmental disorders: Uncomfortable bedfellows. *The Quarterly Journal of Experimental Psychology: Section A, 50*(4), 899–923. https://doi.org/10.1080/713755740

Blomgren, M., & Goberman, A. (2008). Revisiting speech rate and utterance length manipulations in stuttering speakers. *Journal of Communication Disorders, 41*(2), 159–178. https://doi.org/10.1016/j.jcomdis.2007.10.001

Blood, G. (1985). Laterality differences in child stutterers: Heterogeneity, severity levels, and statistical treatments. *Journal of Speech and Hearing Disorders, 50*(1), 66–72. https://doi.org/10.1044/jshd.5001.66

Blood, G., Ridenour Jr., Qualls, C., & Hammer, C. (2003). Co-occurring disorders in children who stutter. *Journal of Communication Disorders, 36*(6), 427–448. https://doi.org/10.1016/S0021-9924(03)00023-6

Bloodstein, O. (1995). *A handbook on stuttering* (5th ed.). Singular.

Bloodstein, O., & Ratner, N. B. (2008). *A handbook on stuttering* (6th ed.). Delmar Learning.

Bloodstein, O., Ratner, N. B., & Brundage, S. B. (2021). *A handbook of stuttering* (7th ed.). Plural Publishing.

Boersma, P., & Weenink, D. (2024). *Praat: Doing phonetics by computer* (Version 6.4.10) [Computer program]. https://www.praat.org/

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

Borden, G. (1983). Initiation versus execution time during manual and oral counting by stutterers. *Journal of Speech and Hearing Research, 26*(3), 389–396. https://doi.org/10.1044/jshr.2603.389

Bosshardt, H. (2006). Cognitive processing load as a determinant of stuttering: Summary of a research programme. *Clinical Linguistics & Phonetics, 20*(5), 371–385. https://doi.org/10.1080/02699200500074321

Bosshardt, H. G. (1999). Effects of concurrent mental calculation on stuttering, inhalation and speech timing. *Journal of Fluency Disorders, 24*(1), 43–72. https://doi.org/10.1016/S0094-730X(98)00022-9

Bosshardt, H. G., & Fransen, H. (1996). Online sentence processing in adults who stutter and adults who do not stutter. *Journal of Speech, Language, and Hearing Research, 39*(4), 785–797. https://doi.org/10.1044/jshr.3904.785

Bothe, A. K., Davidow, J. H., Bramlett, R. E., Franic, D. M., & Ingham, R. J. (2006). Stuttering treatment research 1970–2005: II. Systematic review incorporating trial quality assessment of pharmacological approaches. *American Journal of Speech-Language Pathology*, *15*(4), 342-352. https://doi.org/10.1044/1058-0360(2006/032

Bradshaw, A., Lametti, D., & McGettigan, C. (2021). The role of sensory feedback in developmental stuttering: A review. *Neurobiology of Language, 2*(2), 308–334. DOI: 10.1162/nol_a_00036

Brady, J. (1991). The pharmacology of stuttering: A critical review. *American Journal of Psychiatry, 148*(10), 1309–1316. https://doi.org/10.1176/ajp.148.10.1309

Broadbent, D. (1954). The role of auditory localization in attention and memory span. *Journal of Experimental Psychology, 47*(3), 191–196. https://doi.org/10.1037/h0054182

Brocklehurst, P., & Corley, M. (2011). Investigating the inner speech of people who stutter: Evidence for (and against) the covert repair hypothesis. *Journal of Communication Disorders, 44*(2), 246–260. https://doi.org/10.1016/j.jcomdis.2010.11.004

Brocklehurst, P., Lickley, R., & Corley, M. (2013). Revisiting Bloodstein's anticipatory struggle hypothesis from a psycholinguistic perspective: A variable release threshold hypothesis of stuttering. *Journal of Communication Disorders, 46*(3), 217–237. https://doi.org/10.1016/j.jcomdis.2013.04.002

Brosch, S., & Pirsig, W. (2001). Stuttering in history and culture. *International Journal of Pediatric Otorhinolaryngology, 59*(2), 81–87. https://doi.org/10.1016/S0165-5876(01)00474-8

Brown, B. (2015). Neurocorrelates of speech-motor planning and execution in adults and children who stutter [Doctoral dissertation, University of Iowa]. Iowa Research Online.

Brown, C., Zimmermann, G., Linville, R., & Hegmann, J. (1990). Variations in self-paced behaviors in stutterers and nonstutterers. *Journal of Speech and Hearing Research, 33*(2), 317–323. https://doi.org/10.1044/jshr.3302.317

Buhr, A., & Zebrowski, P. (2009). Sentence position and syntactic complexity of stuttering in early childhood: A longitudinal study. *Journal of Fluency Disorders, 34*(3), 155–172. https://doi.org/10.1016/j.jfludis.2009.08.001

Byrd, C., McGill, M., & Usler, E. (2015). Nonword repetition and phoneme elision in adults who do and do not stutter: Vocal versus nonvocal performance differences. *Journal of Fluency Disorders, 44*, 17–31. https://doi.org/10.1016/j.jfludis.2015.01.004

Byrd, C., Vallely, M., Anderson, J., & Sussman, H. (2012). Nonword repetition and phoneme elision in adults who do and do not stutter. *Journal of Fluency Disorders, 37*(3), 188–201. https://doi.org/10.1016/j.jfludis.2012.03.003

Byrd, D., & Saltzman, E. (2003). The elastic phrase: Modeling the dynamics of boundary-adjacent lengthening. *Journal of Phonetics, 31*(2), 149–180. https://doi.org/10.1016/S0095-4470(02)00085-2

Cai, S., Beal, D., Ghosh, S., Guenther, F., & Perkell, J. (2014). Impaired timing adjustments in response to time-varying auditory perturbation during connected speech production in persons who stutter. *Brain and Language, 129*(1), 24–29. https://doi.org/10.1016/j.bandl.2014.01.002

Cai, S., Beal, D., Ghosh, S., Tiede, M., Guenther, F., & Perkell, J. (2012). Weak responses to auditory feedback perturbation during articulation in persons who stutter: Evidence for abnormal auditory-motor transformation. *PLoS ONE, 7*(7), e41830. https://doi.org/10.1371/journal.pone.0041830

Callan, D., Kent, R., Guenther, F., & Vorperian, H. (2000). An auditory-feedback-based neural network model of speech production that is robust to developmental changes in the size and shape of the articulatory system. *Journal of Speech, Language, and Hearing Research, 43*(3), 721–736. https://doi.org/10.1044/jslhr.4303.721

Caramazza, A. (1984). The logic of neuropsychological research and the problem of patient classification in aphasia. *Brain and Language, 21*(1), 9–20. https://doi.org/10.1016/0093-934X(84)90032-4

Caramazza, A. (1986). On drawing inferences about the structure of normal cognitive systems from the analysis of patterns of impaired performance: The case for single-patient studies. *Brain and Cognition, 5*(1), 41–66. https://doi.org/10.1016/0278-2626(86)90061-8

Caramazza, A. (1997). How many levels of processing are there in lexical access? *Cognitive Neuropsychology, 14*(1), 177–208. https://doi.org/10.1080/026432997381664

Caramazza, A. (2000). Aspects of lexical access: Evidence from aphasia. In A. Grodzinsky, L. Shapiro, & D. Swinney (Eds.), *Language and the brain* (pp. 203–228). Academic Press.

Caramazza, A., & Coltheart, M. (2006). Cognitive neuropsychology twenty years on. *Cognitive Neuropsychology, 23*(1), 3–12. https://doi.org/10.1080/02643290500443250

Caramazza, A., & Hillis, A. (1990). Where do semantic errors come from? *Cortex, 26*(1), 95–122. https://doi.org/10.1016/S0010-9452(13)80077-9

Carpenter, M., & Sommers, R. (1987). Unisensory and bisensory perceptual and memory processing in stuttering adults and normal speakers. *Journal of Fluency Disorders, 12*(4), 291–304. https://doi.org/10.1016/0094-730X(87)90007-6

Carruthers, S., Van Rheenen, T., Gurvich, C., Sumner, P., & Rossell, S. (2019). Characterising the structure of cognitive heterogeneity in schizophrenia spectrum disorders: A systematic review and narrative synthesis. *Neuroscience & Biobehavioral Reviews, 107*, 252–278. https://doi.org/10.1016/j.neubiorev.2019.09.006

Caruso, A., Abbs, J., & Gracco, V. (1988). Kinematic analysis of multiple movement coordination during speech in stutterers. *Brain , 111*(2), 439–455. https://doi.org/10.1093/brain/111.2.439

Chang, S.-E., Below, J., Chow, H., Guenther, F., Hampton Wray, A., Jackson, E., ... Bernstein Ratner, N. (2025). Stuttering: Our current knowledge, research opportunities, and ways to address critical gaps. *Neurobiology of Language, 6*, Article nol_a_00162. https://doi.org/10.1162/nol_a_00162

Chang, S.-E., & Guenther, F. (2020). Involvement of the cortico-basal ganglia-thalamocortical loop in developmental stuttering. *Frontiers in Psychology, 10*, Article 489833. https://doi.org/10.3389/fpsyg.2019.03088

Chang, S.-E., Chow, H., Wieland, E., & McAuley, J. (2016). Relation between functional connectivity and rhythm discrimination in children who do and do not stutter. *NeuroImage: Clinical, 12*, 442–450. https://doi.org/10.1016/j.nicl.2016.08.021

Chen, F. (2011). The relative importance of temporal envelope information for intelligibility prediction: A study on cochlear-implant vocoded speech. *Medical engineering & physics, 33*(8), 1033-1038. https://doi.org/10.1016/j.medengphy.2011.04.004

Choo, A., Smith, S., & Li, H. (2020). Associations between stuttering, comorbid conditions and executive function in children: A population-based study. *BMC Psychology, 8*(1), Article 113. DOI: 10.1186/s40359-020-00481-7

Chowsilpa, S., Bamiou, D., & Koohi, N. (2021). Effectiveness of the auditory temporal ordering and resolution tests to detect central auditory processing disorder in adults with evidence of brain pathology: A systematic review and meta-analysis. *Frontiers in Neurology, 12*, Article 757254. https://doi.org/10.3389/fneur.2021.656117

Ciabarra, A., Elkind, M., Roberts, J., & Marshall, R. (2000). Subcortical infarction resulting in acquired stuttering. *Journal of Neurology, Neurosurgery, and Psychiatry, 69*(4), 546–549. https://doi.org/10.1136/jnnp.69.4.546

Civier, O., Bullock, D., Max, L., & Guenther, F. (2013). Computational modeling of stuttering caused by impairments in a basal ganglia thalamo-cortical circuit involved in syllable selection and initiation. *Brain and Language, 126*(3), 263–278. https://doi.org/10.1016/j.bandl.2013.05.016

Civier, O., Tasko, S., & Guenther, F. (2010). Overreliance on auditory feedback may lead to sound/syllable repetitions: Simulations of stuttering and fluency-inducing conditions with a neural model of speech production. *Journal of Fluency Disorders, 35*(3), 246–279. https://doi.org/10.1016/j.jfludis.2010.05.002

Class, L. (1956). Comparative study of normal speakers and speech defectives with regard to the tactual-kinesthetic perception of form with the tongue [Doctoral dissertation, Ohio State University].

Cohen, L. G., & Starr, A. (1985). Vibration and muscle contraction affect somatosensory evoked potentials. *Neurology*, *35*(5), 691-698. https://doi.org/10.1212/wnl.35.5.691

Coltheart, M. (2017). The assumptions of cognitive neuropsychology: Reflections on Caramazza (1984, 1986). *Cognitive Neuropsychology, 34*(7–8), 397–402. https://doi.org/10.1080/02643294.2017.1324950

Constantino, C., Leslie, P., Quesal, R., & Yaruss, J. (2016). A preliminary investigation of daily variability of stuttering in adults. *Journal of Communication Disorders, 60*, 39–50. https://doi.org/10.1016/j.jcomdis.2016.02.001

Cook, S., Rieger, M., Donlan, C., & Howell, P. (2011). Testing orofacial abilities of children who stutter. *Journal of Fluency Disorders, 36*(1), 27–40. https://doi.org/10.1016/j.jfludis.2011.01.001

Cooper, M., & Allen, G. (1977). Timing control accuracy in normal speakers and stutterers. *Journal of Speech and Hearing Research, 20*(1), 55–71. https://doi.org/10.1044/jshr.2001.55

Corsi, P. M. (1972). *Human memory and the medial temporal region of the brain* [Doctoral dissertation, McGill University].

Cox, D. (2020, September 22). Why some people suffer from a stutter. *BBC Future*. https://www.bbc.com/future/article/20200922-why-some-people-suffer-from-a-stutter

Craig-McQuaide, A., Akram, H., Zrinzo, L., & Tripoliti, E. (2014). A review of brain circuitries involved in stuttering. *Frontiers in Human Neuroscience, 8*, Article 884. https://doi.org/10.3389/fnhum.2014.00884

Crawford, J., & Garthwaite, P. (2007). Comparison of a single case to a control or normative sample in neuropsychology: Development of a Bayesian approach. *Cognitive Neuropsychology, 24*(4), 343–372. https://doi.org/10.1080/02643290701290146

Crawford, J., Garthwaite, P., & Gray, C. D. (2003). Fully operational definitions of dissociations in single-case studies. *Cortex, 39*(2), 357–370. DOI: 10.1016/s0010-9452(08)70117-5

Crawford, J., & Garthwaite, P. (2005). Testing for suspected impairments and dissociations in single-case studies in neuropsychology: Evaluation of alternatives using Monte Carlo simulations and revised tests for dissociations. *Neuropsychology, 19*(3), 318–331. https://doi.org/10.1037/0894-4105.19.3.318

Crawford, J., & Garthwaite, P. (2006). Methods of testing for a deficit in single-case studies: Evaluation of statistical power by Monte Carlo simulation. *Cognitive Neuropsychology, 23*(6), 877–904. https://doi.org/10.1080/02643290500538372

Crawford, J., & Howell, D. (1998). Comparing an individual's test score against norms derived from small samples. *The Clinical Neuropsychologist, 12*(4), 482–486. https://doi.org/10.1076/clin.12.4.482.7241

Crawford, J., Garthwaite, P., Azzalini, A., Howell, D., & Laws, K. (2006). Testing for a deficit in single-case studies: Effects of departures from normality. *Neuropsychologia, 44*(4), 666–677. https://doi.org/10.1016/j.neuropsychologia.2005.06.001

Cummins, F. (2009). Rhythm as entrainment: The case of synchronous speech. *Journal of Phonetics, 37*(1), 16–28. https://doi.org/10.1016/j.wocn.2008.08.003

Cummins, F., & Port, R. (1998). Rhythmic constraints on stress timing in English. *Journal of Phonetics, 26*(2), 145–171. https://doi.org/10.1006/jpho.1998.0070

Curry, F., & Gregory, H. (1969). The performance of stutterers on dichotic listening tasks thought to reflect cerebral dominance. *Journal of Speech and Hearing Research, 12*(1), 73–82. https://doi.org/10.1044/jshr.1201.73

Daliri, A., Prokopenko, R., & Max, L. (2013). Afferent and efferent aspects of mandibular sensorimotor control in adults who stutter. *Journal of Speech, Language, and Hearing Research, 56*(6), 1774–1786. https://doi.org/10.1044/1092-4388(2013/12-0134)

Dalmaijer, E., Nord, C., & Astle, D. (2022). Statistical power for cluster analysis. *BMC bioinformatics*, *23*(1), 205. https://doi.org/10.1186/s12859-022-04675-1

Damian, M. F., Vigliocco, G., & Levelt, W. J. M. (2001). Effects of semantic context in the naming of pictures and words. *Cognition, 81*(3), B77–B86. https://doi.org/10.1016/S0010-0277(01)00135-4

Danks, D., & Davis, I. (2023). Causal inference in cognitive neuroscience. *Wiley Interdisciplinary Reviews: Cognitive Science, 14*(5), e1650. https://doi.org/10.1002/wcs.1650

Darainy, M., Manning, T., & Ostry, D. (2023). Disruption of somatosensory cortex impairs motor learning and retention. *Journal of Neurophysiology, 130*(6), 1521–1528. https://doi.org/10.1152/jn.00231.2023

Davidson, A., Eitel, M., Lange, R., French, L., Lippa, S., Brickell, T., & Brungart, D. (2023). Efficient estimation of the binaural masking level difference using a technique based on manual audiometry. *Journal of Speech, Language, and Hearing Research, 66*(4), 1378–1393. https://doi.org/10.1044/2022_JSLHR-22-00519

de Jong, N., & Wempe, T. (2009). Praat script to detect syllable nuclei and measure speech rate automatically. *Behavior Research Methods, 41*(2), 385–390. https://doi.org/10.3758/BRM.41.2.385

De Nil, L., & Abbs, J. (1991). Kinaesthetic acuity of stutterers and non-stutterers for oral and non-oral movements. *Brain, 114*(5), 2145–2158. https://doi.org/10.1093/brain/114.5.2145

de Saint-Exupéry, A. (2002). *Le petit prince* (Éd. spéciale). Gallimard Jeunesse (Folio Junior).

De Visscher, A., & Noël, M.-P. (2013). A case study of arithmetic facts dyscalculia caused by a hypersensitivity-to-interference in memory. *Cortex, 49*(1), 50–70. https://doi.org/10.1016/j.cortex.2012.01.003

Demanez, L., Dony-Closon, B., Lhonneux-Ledoux, E., & Demanez, J. (2003). Central auditory processing assessment: A French-speaking battery. *Acta Oto-Rhino-Laryngologica Belgica, 57*(4), 275–290.

Dilley, L., Wallace, J., & Heffner, C. (2012). Perceptual isochrony and fluency in speech by normal talkers under varying task demands. In O. F. E. Bohn & M. J. Munro (Eds.), *Understanding prosody* (pp. 237–258). De Gruyter Mouton.

Doherty, M., Buchy, P., Standaert, B., Giaquinto, C., & Prado-Cohrs, D. (2016). Vaccine impact: Benefits for human health. *Vaccine, 34*(52), 6707–6714. https://doi.org/10.1016/j.vaccine.2016.10.025

Eberhardt, F., & Scheines, R. (2007). Interventions and causal inference. *Philosophy of Science, 74*(5), 981–995. https://doi.org/10.1086/525638

Ebrahimi, S., & Ostry, D. (2024). The human somatosensory cortex contributes to the encoding of newly learned movements. *Proceedings of the National Academy of Sciences, 121*(6), e2316294121. https://doi.org/10.1073/pnas.2316294121

Eggers, K., De Nil, L., & Van den Bergh, B. (2010). Temperament dimensions in stuttering and typically developing children. *Journal of Fluency Disorders, 35*(4), 355–372. https://doi.org/10.1016/j.jfludis.2010.10.004

Eggers, K., De Nil, L., & Van den Bergh, B. (2013). Inhibitory control in childhood stuttering. *Journal of Fluency Disorders, 38*(1), 1–13. https://doi.org/10.1016/j.jfludis.2012.10.001

Eichorn, N., Hall, J., & Marton, K. (2023). Complex working memory in adults with and without stuttering disorders: Performance patterns and predictive relationships. *Journal of Fluency Disorders, 77*, Article 105993. https://doi.org/10.1016/j.jfludis.2023.105993

Eichorn, N., Marton, K., Schwartz, R., Melara, R., & Pirutinsky, S. (2016). Does working memory enhance or interfere with speech fluency in adults who do and do not stutter? Evidence from a dual-task paradigm. *Journal of Speech, Language, and Hearing Research, 59*(3), 415-429. https://doi.org/10.1044/2015_JSLHR-S-15-0249

Eichorn, N., Pirutinsky, S., & Marton, K. (2019). Effects of different attention tasks on concurrent speech in adults who stutter and fluent controls. *Journal of Fluency Disorders, 61*, Article 105714. https://doi.org/10.1016/j.jfludis.2019.105714

Ekhart, C., Van Hunsel, F., Van Harten, P., Van Baarsen, J., Yingying, T., & Bast, B. (2021). Drug-induced stuttering: Occurrence and possible pathways. *Frontiers in Psychiatry, 12*, Article 692568. https://doi.org/10.3389/fpsyt.2021.692568

Ellis, A., & Young, A. (1988). *Human cognitive neuropsychology*. Lawrence Erlbaum Associates.

Eriksen, B., & Eriksen, C. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics, 16*(1), 143–149. https://doi.org/10.3758/BF03203267

Eronen, M. (2020). Causal discovery and the problem of psychological interventions. *New Ideas in Psychology, 59*, Article 100785. https://doi.org/10.1016/j.newideapsych.2020.100785

Etchell, A., Civier, O., Ballard, K., & Sowman, P. (2018). A systematic literature review of neuroimaging research on developmental stuttering between 1995 and 2016. *Journal of Fluency Disorders, 55*, 6–45. https://doi.org/10.1016/j.jfludis.2017.03.007

Etchell, A., Johnson, B., & Sowman, P. (2014). Behavioral and multimodal neuroimaging evidence for a deficit in brain timing networks in stuttering: A hypothesis and theory. *Frontiers in Human Neuroscience, 8*, Article 467. https://doi.org/10.3389/fnhum.2014.00467

Falk, S., Maslow, E., Thum, G., & Hoole, P. (2016). Temporal variability in sung productions of adolescents who stutter. *Journal of Communication Disorders, 62*, 101–114. https://doi.org/10.1016/j.jcomdis.2016.05.012

Falk, S., Müller, T., & Bella, S. (2014). Sensorimotor synchronization in stuttering children and adolescents. *Procedia - Social and Behavioral Sciences, 126*, 206–207. DOI: 10.1016/j.sbspro.2014.02.375

Falk, S., Müller, T., & Dalla Bella, S. (2015). Non-verbal sensorimotor timing deficits in children and adolescents who stutter. *Frontiers in Psychology, 6*, Article 847. https://doi.org/10.3389/fpsyg.2015.00847

Farazi, M., Dastgerdi, Z., Lotfi, Y., Moossavi, A., & Bakhshi, E. (2023). Effect of an Auditory Temporal Training Program on Speech Fluency of Children with Developmental Stuttering. *Iranian Journal of Child Neurology, 17*(1), 39–53. https://doi.org/10.22037/ijcn.v17i1.35885

Farrell, S., & Lewandowsky, S. (2018). *Computational modeling of cognition and behavior*. Cambridge University Press.

Fedorenko, E. (2021). The early origins and the growing popularity of the individual-subject analytic approach in human neuroscience. *Current Opinion in Behavioral Sciences, 40*, 105–112. https://doi.org/10.1016/j.cobeha.2021.02.023

Feinberg, A., Griffin, B., & Levey, M. (2000). Psychological aspects of chronic tonic and clonic stuttering: Suggested therapeutic approaches. *American Journal of Orthopsychiatry, 70*(4), 465–473. https://psycnet.apa.org/doi/10.1037/h0087709

Fodor, J. A. (1983). *The modularity of mind*. MIT Press.

Folkins, J., & Abbs, J. (1975). Lip and jaw motor control during speech: Responses to resistive loading of the jaw. *Journal of Speech and Hearing Research, 18*(1), 207–220. https://doi.org/10.1044/jshr.1801.207

Foundas, A., Bollich, A., Feldman, J., Corey, D., Hurley, M., Lemen, L., & Heilman, K. (2004). Aberrant auditory processing and atypical planum temporale in developmental stuttering. *Neurology, 63*(9), 1640–1646. https://doi.org/10.1212/01.WNL.0000142993.33158.2A

Foundas, A., Mock, J., Corey, D., Golob, E., & Conture, E. (2013). The SpeechEasy device in stuttering and nonstuttering adults: fluency effects while speaking and reading. *Brain and Language, 126*(2), 141–150. https://doi.org/10.1016/j.bandl.2013.04.004

Fox, N., Leonard, M., Sjerps, M., & Chang, E. (2020). Transformation of a temporal speech cue to a spatial neural code in human auditory cortex. *eLife, 9*, Article e53051. https://doi.org/10.7554/eLife.53051

Frankford, S., Cai, S., Nieto-Castañón, A., & Guenther, F. (2022). Auditory feedback control in adults who stutter during metronome-paced speech II. Formant perturbation. *Journal of Fluency Disorders, 74*, Article 105928. https://doi.org/10.1016/j.jfludis.2022.105928

Frith, U. (2001). What framework should we use for understanding developmental disorders? *Developmental Neuropsychology, 20*(2), 555-563. https://doi.org/10.1207/S15326942DN2002_6

Frith, U., & Kelly, S. W. (1999). *Neuropsychological case studies of developmental disorders*. Psychology Press.

Fruhstorfer, H., Gross, W., & Selbmann, O. (2001). Von Frey hairs: New materials for a new design. *European Journal of Pain, 5*(3), 341–342. https://doi.org/10.1053/eujp.2001.0250

Fucci, D., Petrosino, L., Gorman, P., & Harris, D. (1985). Vibrotactile magnitude production scaling: A method for studying sensory-perceptual responses of stutterers and fluent speakers. *Journal of Fluency Disorders, 10*(1), 69–75. https://doi.org/10.1016/0094-730X(85)90007-5

Fucci, D., Petrosino, L., Schuster, S., & Belch, M. (1991). Lingual vibrotactile threshold shift differences between stutterers and normal speakers during magnitude-estimation scaling. *Perceptual and Motor Skills, 73*(1), 55–62. https://doi.org/10.2466/pms.1991.73.1.55

Garnett, E., McAuley, J., Wieland, E., Chow, H., Zhu, D., Dilley, L., & Chang, S.-E. (2023). Auditory rhythm discrimination in adults who stutter: An fMRI study. *Brain and Language, 236*, Article 105219. https://doi.org/10.1016/j.bandl.2022.105219

Gastaldon, S., Bonfiglio, N., Vespignani, F., & Peressotti, F. (2024). Predictive language processing: Integrating comprehension and production, and what atypical populations can tell us. *Frontiers in Psychology, 15*, Article 1369177. https://doi.org/10.3389/fpsyg.2024.1369177

Gastaldon, S., Busan, P., Molinaro, N., & Lizarazu, M. (2024). Cortical tracking of speech is reduced in adults who stutter when listening for speaking. *Journal of Speech, Language, and Hearing Research, 67*(11), 4339-4357. https://doi.org/10.1044/2024_JSLHR-24-00227

Gattie, M., Lieven, E., & Kluk, K. (2025). Adult stuttering prevalence II: Recalculation, subgrouping and estimate of stuttering community engagement. *Journal of Fluency Disorders, 83*, Article 106086. https://doi.org/10.1016/j.jfludis.2024.106086

Gervain, J. (2018). The role of prenatal experience in language development. *Current Opinion in Behavioral Sciences, 21*, 62–67. https://doi.org/10.1016/j.cobeha.2018.02.004

Gilbert, E., Mérette, C., Jomphe, V., Émond, C., Rouleau, N., Bouchard, R. H., ... Maziade, M. (2014). Cluster analysis of cognitive deficits may mark heterogeneity in schizophrenia in terms of outcome and response to treatment. *European Archives of Psychiatry and Clinical Neuroscience, 264*(4), 333–343. doi: 10.1007/s00406-013-0463-7

Giofrè, D., Toffalini, E., Provazza, S., Calcagnì, A., Altoè, G., & Roberts, D. J. (2019). Are children with developmental dyslexia all the same? A cluster analysis with more than 300 cases. *Dyslexia, 25*(3), 284–295. https://doi.org/10.1002/dys.1629

Glover, H., Kalinowski, J., Rastatter, M., & Stuart, A. (1996). Effect of instruction to sing on stuttering frequency at normal and fast rates. *Perceptual and Motor Skills, 83*(2), 511–522. https://doi.org/10.2466/pms.1996.83.2.511

Gracco, V., & Abbs, J. (1989). Sensorimotor characteristics of speech motor sequences. *Experimental Brain Research, 75*(3), 586–598. DOI: 10.1007/BF00249910

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

Guenther, F. H. (1994). A neural network model of speech acquisition and motor equivalent speech production. *Biological Cybernetics, 72*(1), 43–53. https://doi.org/10.1007/BF00206237

Guenther, F. H., & Hickok, G. (2015). Role of the auditory system in speech production. In S. L. Ackermann & A. R. Crossman (Eds.), *Handbook of clinical neurology* (Vol. 129, pp. 161–175). Elsevier.

Guitar, B. (1998). *Stuttering: An integrated approach to its nature and treatment* (2nd ed.). Williams & Wilkins.

Guitar, B. (2023). *Stuttering: An integrated approach to its nature and treatment* (6th ed.). Lippincott Williams & Wilkins.

Gunderson, D. (1972). *Some relations between oral form discrimination and stuttering* [Unpublished master's thesis]. University of Minnesota.

Haggard, P., & de Boer, L. (2014). Oral somatosensory awareness. *Neuroscience & Biobehavioral Reviews, 47*, 469–484. https://doi.org/10.1016/j.neubiorev.2014.09.015

Hakim, H. B., & Ratner, N. B. (2004). Nonword repetition abilities of children who stutter: An exploratory study. *Journal of Fluency Disorders, 29*(3), 179–199. https://doi.org/10.1016/j.jfludis.2004.06.001

Hall, J., & Jerger, J. (1978). Central auditory function in stutterers. *Journal of Speech and Hearing Research, 21*(2), 324–337. https://doi.org/10.1044/jshr.2102.324

Hampton, A., & Weber-Fox, C. (2008). Non-linguistic auditory processing in stuttering: Evidence from behavior and event-related brain potentials. *Journal of Fluency Disorders, 33*(4), 253–273. https://doi.org/10.1016/j.jfludis.2008.08.001

Harley, T. A. (2013). *The psychology of language: From data to theory* (4th ed.). Psychology Press.

Harris, D., Fucci, D., & Petrosino, L. (1991). Magnitude estimation and cross-modal matching of auditory and lingual vibrotactile sensation by normal speakers and stutterers. *Journal of Speech, Language, and Hearing Research, 34*(1), 177–182. https://doi.org/10.1044/jshr.3401.177

Hawkins, S. (2014). Situational influences on rhythmicity in speech, music, and their interaction. *Philosophical Transactions of the Royal Society B: Biological Sciences, 369*(1658), Article 20130398. https://doi.org/10.1098/rstb.2013.0398

Hennessey, N., Nang, C., & Beilby, J. (2008). Speeded verbal responding in adults who stutter: Are there deficits in linguistic encoding? *Journal of Fluency Disorders, 33*(3), 180–202. https://doi.org/10.1016/j.jfludis.2008.06.001

Hickok, G., Houde, J., & Rong, F. (2011). Sensorimotor integration in speech processing: Computational basis and neural organization. *Neuron, 69*(3), 407–422. https://doi.org/10.1016/j.neuron.2011.01.019

Hilger, A., Zelaznik, H., & Smith, A. (2016). Evidence that bimanual motor timing performance is not a significant factor in developmental stuttering. *Journal of Speech, Language, and Hearing Research, 59*(4), 674–685. doi: 10.1044/2016_JSLHR-S-15-0172

Hillis, A. E., Rapp, B., Romani, C., & Caramazza, A. (1990). Selective impairment of semantics in lexical processing. *Cognitive Neuropsychology, 7*(3), 191–243. https://doi.org/10.1080/02643299008253442

Hirsh, I. J. (1948). The influence of interaural phase on interaural summation and inhibition. *The Journal of the Acoustical Society of America, 20*(4), 536–544. https://doi.org/10.1121/1.1906407

Hirsh, I. J. (1959). Auditory perception of temporal order. *The Journal of the Acoustical Society of America, 31*(6), 759–767. https://doi.org/10.1121/1.1907782

Hoole, P. (1987). Bite-block speech in the absence of oral sensibility. In *Proceedings of the XIth International Congress of Phonetic Sciences* (pp. 16–19). Tallinn.

Houde, J., & Nagarajan, S. (2011). Speech production as state feedback control. *Frontiers in Human Neuroscience, 5*, Article 82. https://doi.org/10.3389/fnhum.2011.00082

Howell, P. (2004). Assessment of some contemporary theories of stuttering that apply to spontaneous speech. *Contemporary Issues in Communication Science and Disorders, 31*, 123–140.

Howell, P. (2007). Signs of developmental stuttering up to age eight and at 12 plus. *Clinical Psychology Review, 27*(3), 287–306. https://doi.org/10.1016/j.cpr.2006.08.005

Howell, P. (2010). Behavioral effects arising from the neural substrates for atypical planning and execution of word production in stuttering. *Experimental Neurology, 225*(1), 55–59. https://doi.org/10.1016/j.expneurol.2010.06.012

Howell, P., & Au-Yeung, J. (2002). The EXPLAN theory of fluency control applied to the diagnosis of stuttering. In E. Fava (Ed.), *Pathology and therapy of speech disorders* (pp. 75–94). John Benjamins.

Howell, P., Au-Yeung, J., & Sackin, S. (2000). Internal structure of content words leading to lifespan differences in phonological difficulty in stuttering. *Journal of Fluency Disorders, 25*(1), 1–20. https://doi.org/10.1016/S0094-730X(99)00025-X

Howell, P., El-Yaniv, N., & Powell, D. J. (1987). Factors affecting fluency in stutterers when speaking under altered auditory feedback. In H. F. M. Peters & W. Hulstijn (Eds.), *Speech motor dynamics in stuttering* (pp. 361–369). Springer.

Howell, P., Sackin, S., & Williams, R. (1999). Differential effects of frequency-shifted feedback between child and adult stutterers. *Journal of Fluency Disorders, 24*(2), 127–136. https://doi.org/10.1016/S0094-730X(98)00021-7

Hulstijn, W., Summers, J., Van Lieshout, P., & Peters, H. (1992). Timing in finger tapping and speech: A comparison between stutterers and fluent speakers. *Human Movement Science, 11*(1–2), 113–124. https://doi.org/10.1016/0167-9457(92)90054-F

Hutchinson, J., & Ringel, R. (1975). The effect of oral sensory deprivation on stuttering behavior. *Journal of Communication Disorders, 8*(3), 249–258. https://doi.org/10.1016/0021-9924(75)90017-9

Hyndman, R., & Fan, Y. (1996). Sample quantiles in statistical packages. *The American Statistician, 50*(4), 361–365. https://doi.org/10.1080/00031305.1996.10473566

Ingham, R., Warner, A., Byrd, A., & Cotton, J. (2006). Speech effort measurement and stuttering: Investigating the chorus reading effect. *Journal of Speech, Language, and Hearing Research, 49*(3), 660–670. https://doi.org/10.1044/1092-4388(2006/048)

Ingham, R., & Carroll, P. (1977). Listener judgment of differences in stutterers' nonstuttered speech during chorus and nonchorus reading conditions. *Journal of Speech and Hearing Research, 20*(2), 293–302. https://doi.org/10.1044/jshr.2002.293

Ito, T., & Gomi, H. (2007). Cutaneous mechanoreceptors contribute to the generation of a cortical reflex in speech. *Neuroreport, 18*(9), 907–910. https://doi.org/10.1097/WNR.0b013e32810f2dfb

Iverach, L., Jones, M., McLellan, L., Lyneham, H., Menzies, R., Onslow, M., & Rapee, R. (2016). Prevalence of anxiety disorders among children who stutter. *Journal of Fluency Disorders, 49*, 13–28. https://doi.org/10.1016/j.jfludis.2016.07.002

Jackson, N., & Coltheart, M. (2013). *Routes to reading success and failure: Toward an integrated cognitive psychology of atypical reading*. Psychology Press.

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2021). *An introduction to statistical learning: With applications in R* (2nd ed.). Springer. https://doi.org/10.1007/978-1-0716-1418-1

Jensen, P., Sheehan, J., Williams, W., & LaPointe, L. (1975). Oral sensory-perceptual integrity of stutterers. *Folia Phoniatrica et Logopaedica, 27*(1), 38–45. https://doi.org/10.1159/000263966

Jonas, E., & Kording, K. (2017). Could a neuroscientist understand a microprocessor? *PLoS Computational Biology, 13*(1), e1005268. https://doi.org/10.1371/journal.pcbi.1005268

Jones, M. (1976). Time, our lost dimension: Toward a new theory of perception, attention, and memory. *Psychological Review, 83*(5), 323–355. https://doi.org/10.1037/0033-295X.83.5.323

Jones, R., Choi, D., Conture, E., & Walden, T. (2014). Temperament, emotion, and childhood stuttering. *Seminars in Speech and Language, 35*(2), 114–131. Doi: 10.1055/s-0034-1371755

Kaddour, J., Lynch, A., Liu, Q., Kusner, M. J., & Silva, R. (2022). Causal machine learning: A survey and open problems. *arXiv Preprint*, arXiv:2206.15475. https://arxiv.org/abs/2206.15475

Kalinowski, J., Armson, J., Stuart, A., Roland-Mieszkowski, M., & Gracco, V. (1993). Effects of alterations in auditory feedback and speech rate on stuttering frequency. *Language and Speech, 36*(1), 1–16. https://doi.org/10.1177/002383099303600101

Kalinowski, J., Stuart, A., Sark, S., & Armson, J. (1996). Stuttering amelioration at various auditory feedback delays and speech rates. *International Journal of Language & Communication Disorders, 31*(3), 259–269. DOI: 10.3109/13682829609033157

Karmiloff-Smith, A. (1994). Précis of Beyond modularity: A developmental perspective on cognitive science. *Behavioral and Brain Sciences, 17*(4), 693–707. https://psycnet.apa.org/doi/10.1017/S0140525X00036621

Karmiloff-Smith, A. (1997). Crucial differences between developmental cognitive neuroscience and adult neuropsychology. *Developmental Neuropsychology, 13*(4), 513–524. https://psycnet.apa.org/doi/10.1080/87565649709540693

Karmiloff-Smith, A. (1998). Development itself is the key to understanding developmental disorders. *Trends in Cognitive Sciences, 2*(10), 389–398. https://doi.org/10.1016/S1364-6613(98)01230-3

Karmiloff-Smith, A. (2012). Foreword: Development is not about studying children: The importance of longitudinal approaches. *American Journal on Intellectual and Developmental Disabilities, 117*(2), 87–89. https://doi.org/10.1352/1944-7558-117.2.87

Kay, J., & Ellis, A. (1987). A cognitive neuropsychological case study of anomia: Implications for psychological models of word retrieval. *Brain, 110*(3), 613–629. https://doi.org/10.1093/brain/110.3.613

Kekade, N., & Valame, D. (2014). Auditory temporal processing in children with stuttering. *Journal of Indian Speech Language & Hearing Association, 28*(2), 41–44.

Kell, C., Neumann, K., von Kriegstein, K., Posenenske, C., von Gudenberg, A., Euler, H., & Giraud, A. (2009). How the brain repairs stuttering. *Brain: A Journal of Neurology, 132*(10), 2747–2760. https://doi.org/10.1093/brain/awp185

Kent, R. D. (2024). The feel of speech: Multisystem and polymodal somatosensation in speech production. *Journal of Speech, Language, and Hearing Research, 67*(5), 1424–1460. DOI: 10.1044/2024_JSLHR-23-00575

Kent, R. D. (1984). Stuttering as a temporal programming disorder. In R. F. Curlee & W. H. Perkins (Eds.), *Nature and treatment of stuttering: New directions* (pp. 97–114). College Hill Press.

Kent, R. D., Kim, Y., & Chen, L. (2022). Oral and laryngeal diadochokinesis across the life span: A scoping review of methods, reference data, and clinical applications. *Journal of Speech, Language, and Hearing Research, 65*(2), 574–623. DOI: 10.1044/2021_JSLHR-21-00396

Kiefte, M., & Armson, J. (2008). Dissecting choral speech: Properties of the accompanist critical to stuttering reduction. *Journal of Communication Disorders, 41*(1), 33–48. https://doi.org/10.1016/j.jcomdis.2007.03.002

Kimura, D. (1961). Cerebral dominance and the perception of verbal stimuli. *Canadian Journal of Psychology, 15*(3), 166–171. https://doi.org/10.1037/h0083219

Kleinow, J., & Smith, A. (2000). Influences of length and syntactic complexity on the speech motor stability of the fluent speech of adults who stutter. *Journal of Speech, Language, and Hearing Research, 43*(2), 548–559. https://doi.org/10.1044/jslhr.4302.548

Koca, T., Belgin, E., & Ölçek, G. (2024). Investigation of central auditory processing performance in individuals with and without stuttering. *Journal of Fluency Disorders, 80*, Article 106048. https://doi.org/10.1016/j.jfludis.2024.106048

Kolk, H., & Postma, A. (1997). Stuttering as a covert-repair phenomenon. In R. F. Curlee & G. M. Siegel (Eds.), *Nature and treatment of stuttering: New directions* (pp. 182–203). Allyn & Bacon.

Korzeczek, A., Cholin, J., Jorschick, A., Hewitt, M., & Sommer, M. (2020). Finger sequence learning in adults who stutter. *Frontiers in Psychology, 11*, Article 1543. https://doi.org/10.3389/fpsyg.2020.01543

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

Krakauer, J., Ghazanfar, A., Gomez-Marin, A., MacIver, M., & Poeppel, D. (2017). Neuroscience needs behavior: Correcting a reductionist bias. *Neuron, 93*(3), 480–490. https://doi.org/10.1016/j.neuron.2016.12.041

Lane, H., & Tranel, B. (1971). The Lombard sign and the role of hearing in speech. *Journal of Speech and Hearing Research, 14*(4), 677–709. https://doi.org/10.1044/jshr.1404.677

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

Lawless, H. T., & Heymann, H. (2010). Measurement of sensory thresholds. In *Sensory evaluation of food: Principles and practices* (pp. 125–147). Springer. https://doi.org/10.1007/978-1-4419-6488-5_7

Lawrence, M., & Barclay, I. (1998). Stuttering: A brief review. *American Family Physician, 57*(9), 2175–2178.

Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. MIT Press.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences, 22*(1), 1–38. https://doi.org/10.1017/S0140525X99001776

Lincoln, M., Packman, A., & Onslow, M. (2006). Altered auditory feedback and the treatment of stuttering: A review. *Journal of Fluency Disorders, 31*(2), 71–89. https://doi.org/10.1016/j.jfludis.2006.04.001

Lincoln, M., Packman, A., Onslow, M., & Jones, M. (2010). An experimental investigation of the effect of altered auditory feedback on the conversational speech of adults who stutter. *Journal of Speech, Language, and Hearing Research, 53*(5), 1122–1131. https://doi.org/10.1044/1092-4388(2009/07-0266)

Logan, K., Roberts, R., Pretto, A., & Morey, M. (2002). Speaking slowly: Effects of four self-guided training approaches on adults' speech rate and naturalness. *American Journal of Speech-Language Pathology, 11*(2), 163–174. https://doi.org/10.1044/1058-0360(2002/016)

Lotfi, Y., Dastgerdi, Z., Farazi, M., Moossavi, A., & Bakhshi, E. (2020). Auditory temporal processing assessment in children with developmental stuttering. *International Journal of Pediatric Otorhinolaryngology, 132*, Article 109935. https://doi.org/10.1016/j.ijporl.2020.109935

Loucks, T., & De Nil, L. F. (2006b). Oral kinesthetic deficit in adults who stutter: A target-accuracy study. *Journal of Motor Behavior, 38*(3), 238–246. DOI: 10.3200/JMBR.38.3.238-247

Loucks, T., & De Nil, L. F. (2006a). Anomalous sensorimotor integration in adults who stutter: A tendon vibration study. *Neuroscience Letters, 402*(1–2), 195–200. https://doi.org/10.1016/j.neulet.2006.04.002

Loucks, T., & De Nil, L. F. (2012). Oral sensorimotor integration in adults who stutter. *Folia Phoniatrica et Logopaedica, 64*(3), 116–121. https://doi.org/10.1159/000338248

Loucks, T., De Nil, L. F., & Sasisekaran, J. (2007). Jaw-phonatory coordination in chronic developmental stuttering. *Journal of Communication Disorders, 40*(3), 257–272. https://doi.org/10.1016/j.jcomdis.2006.06.016

Loucks, T., Shosted, R., De Nil, L. F., Poletto, C., & King, A. (2010). Coordinating voicing onset with articulation: A potential role for sensory cues in shaping phonological distinctions. *Phonetica, 67*(1–2), 47–62. DOI: 10.1159/000319378

Lundgren, K., Helm-Estabrooks, N., & Klein, R. (2010). Stuttering following acquired brain damage: A review of the literature. *Journal of Neurolinguistics, 23*(5), 447–454. https://doi.org/10.1016/j.jneuroling.2009.08.008

Macleod, J., Kalinowski, J., Stuart, A., & Armson, J. (1995). Effect of single and combined altered auditory feedback on stuttering frequency at two speech rates. *Journal of Communication Disorders, 28*(3), 217–228. https://doi.org/10.1016/0021-9924(94)00010-W

MacNeilage, P., Rootes, T., & Chase, R. (1967). Speech production and perception in a patient with severe impairment of somesthetic perception and motor control. *Journal of Speech, Language, and Hearing Research, 10*(3), 449–467.

Maguire, G. A., Nguyen, D. L., Simonson, K. C., & Kurz, T. L. (2020). The pharmacologic treatment of stuttering and its neuropharmacologic basis. *Frontiers in Neuroscience, 14*, Article 158. https://doi.org/10.3389/fnins.2020.00158

Maguire, G., Riley, G., Franklin, D., Maguire, M., Nguyen, C., & Brojeni, P. (2004). Olanzapine in the treatment of developmental stuttering: A double-blind, placebo-controlled trial. *Annals of Clinical Psychiatry, 16*(2), 63–67. DOI: 10.1080/10401230490452834

Mahon, B., Costa, A., Peterson, R., Vargas, K., & Caramazza, A. (2007). Lexical selection is not by competition: A reinterpretation of semantic interference and facilitation effects in the picture-word interference paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*(3), 503–535. https://doi.org/10.1037/0278-7393.33.3.503

Majeed, N. M., Chua, Y. J., Kothari, M., Kaur, M., Quek, F. Y., Ng, M. H., ... Hartanto, A. (2023). Anxiety disorders and executive functions: A three-level meta-analysis of

reaction time and accuracy. *Psychiatry Research Communications, 3*(1), Article 100100. DOI: 10.1080/10401230490452834

Marslen-Wilson, W., & Tyler, L. (1981). Central processes in speech understanding. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences, 295*, 317–332. https://doi.org/10.1098/rstb.1981.0143

Martin, J. (1972). Rhythmic (hierarchical) versus serial structure in speech and other behavior. *Psychological Review, 79*(6), 487–509. DOI: 10.1037/h0033467

Martin, R., Lawrence, B., Haroldson, S., & Gunderson, D. (1981). Stuttering and oral stereognosis. *Perceptual and Motor Skills, 53*(1), 155–162. https://doi.org/10.2466/pms.1981.53.1.155

Masapollo, M., Segawa, J. A., Beal, D. S., Tourville, J. A., Nieto-Castañón, A., Heyne, M., ... Guenther, F. H. (2021). Behavioral and neural correlates of speech motor sequence learning in stuttering and neurotypical speakers: An fMRI investigation. *Neurobiology of Language, 2*(1), 106–137. DOI: 10.1162/nol_a_00027

Matsuhashi, K., Itahashi, T., Aoki, R., & Hashimoto, R. (2023). Meta-analysis of structural integrity of white matter and functional connectivity in developmental stuttering. *Brain Research Bulletin, 205*, Article 110827. https://doi.org/10.1016/j.brainresbull.2023.110827

Max, L., & Daliri, A. (2019). Limited pre-speech auditory modulation in individuals who stutter: Data and hypotheses. *Journal of Speech, Language, and Hearing Research, 62*(8S), 3071–3084. https://doi.org/10.1044/2019_JSLHR-S-CSMC7-18-0358

Max, L., & Gracco, V. (2005). Coordination of oral and laryngeal movements in the perceptually fluent speech of adults who stutter. *Journal of Speech, Language, and Hearing Research, 48*(3), 524–542. https://doi.org/10.1044/1092-4388(2005/036)

Max, L., & Yudman, E. M. (2003). Accuracy and variability of isochronous rhythmic timing across motor systems in stuttering versus nonstuttering individuals. *Journal of Speech, Language, and Hearing Research, 46*(1), 146–163. https://doi.org/10.1044/1092-4388(2003/012)

Max, L., Caruso, A., & Gracco, V. (2003). Kinematic analyses of speech, orofacial nonspeech, and finger movements in stuttering and nonstuttering adults. *Journal of Speech, Language, and Hearing Research, 46*(1), 215–232. DOI: 10.1044/1092-4388(2003/017)

Max, L., Guenther, F., & Gracco, V. (2004). Unstable or insufficiently activated internal models and feedback-biased motor control as sources of dysfluency: A theoretical model of stuttering. *Contemporary Issues in Communication Science and Disorders, 31*(Spring), 105–122.

Maxfield, N. (2020). Inhibitory control of lexical selection in adults who stutter. *Journal of Fluency Disorders, 66*, Article 105780. https://doi.org/10.1016/j.jfludis.2020.105780

Maxfield, N. (2021). Cognitive control of action naming in adults who stutter. *Journal of Fluency Disorders, 70*, Article 105841. https://doi.org/10.1016/j.jfludis.2021.105841

Maxfield, N. D. (2017). Semantic and phonological encoding times in adults who stutter: Brain electrophysiological evidence. *Journal of Speech, Language, and Hearing Research, 60*(10), 2906–2923. doi: 10.1044/2017_JSLHR-L-16-0309

Maxfield, N., Olsen, W., Kleinman, D., Frisch, S. A., Ferreira, V., & Lister, J. (2016). Attention demands of language production in adults who stutter. *Clinical Neurophysiology, 127*(4), 1942–1960. doi: 10.1016/j.clinph.2016.01.016

McAuley, J., & Jones, M. (2003). Modeling effects of rhythmic context on perceived duration: A comparison of interval and entrainment approaches to short-interval timing. *Journal*

of Experimental Psychology: Human Perception and Performance, 29(6), 1102–1125. https://doi.org/10.1037/0096-1523.29.6.1102

McCloskey, M., & Caramazza, A. (1988). Theory and methodology in cognitive neuropsychology: A response to our critics. *Cognitive Neuropsychology, 5*(5), 583–623. https://doi.org/10.1080/02643298808253276

McKnight, R., & Cullinan, W. (1987). Subgroups of stuttering children: Speech and voice reaction times, segmental durations, and naming latencies. *Journal of Fluency Disorders, 12*(3), 217–233. https://doi.org/10.1016/0094-730X(87)90028-3

Meier, A. M., & Guenther, F. H. (2023). Neurocomputational modeling of speech motor development. *Journal of Child Language, 50*(6), 1318–1335. DOI: 10.1017/S0305000923000260

Moore, B. C. (2003). Temporal integration and context effects in hearing. *Journal of Phonetics, 31*(3–4), 563–574. https://doi.org/10.1016/S0095-4470(03)00011-1

Morton, J., & Frith, U. (1995). Causal modeling: A structural approach to developmental psychopathology. In D. Cicchetti & D. J. Cohen (Eds.), *Developmental psychopathology* (Vol. 1, pp. 357–390). Wiley.

Musiek, F., & Weihing, J. (2011). Perspectives on dichotic listening and the corpus callosum. *Brain and Cognition, 76*(2), 225–232. https://doi.org/10.1016/j.bandc.2011.03.011

Musiek, F., Baran, J., & Pinheiro, M. (1990). Duration pattern recognition in normal subjects and patients with cerebral and cochlear lesions. *Audiology: Official Organ of the International Society of Audiology, 29*(6), 304–313. DOI: 10.3109/00206099009072861

Mysak, E. (1960). Servo theory and stuttering. *The Journal of Speech and Hearing Disorders, 25*, 188–195.

Namasivayam, A. K., & Van Lieshout, P. (2008). Investigating speech motor practice and learning in people who stutter. *Journal of Fluency Disorders, 33*(1), 32–51. https://doi.org/10.1016/j.jfludis.2007.11.005

Namasivayam, A. K., & Van Lieshout, P. (2011). Speech motor skill and stuttering. *Journal of Motor Behavior, 43*(6), 477–489. https://doi.org/10.1080/00222895.2011.628347

Navarrete, E., Basagni, B., Alario, F., & Costa, A. (2006). Does word frequency affect lexical selection in speech production? *Quarterly Journal of Experimental Psychology, 59*(10), 1681–1690. DOI: 10.1080/17470210600750558

Navarrete, E., Scaltritti, M., Mulatti, C., & Peressotti, F. (2013). Age-of-acquisition effects in delayed picture-naming tasks. *Psychonomic Bulletin & Review, 20*(1), 148–153. DOI: 10.3758/s13423-012-0310-2

Neef, N. E., Jung, K., Rothkegel, H., Pollok, B., von Gudenberg, A. W., Paulus, W., & Sommer, M. (2011). Right-shift for non-speech motor processing in adults who stutter. *Cortex, 47*(8), 945–954. https://doi.org/10.1016/j.cortex.2010.06.007

Neilson, M. D., & Neilson, P. D. (1987). Speech motor control and stuttering: A computational model of adaptive sensory-motor processing. *Speech Communication, 6*(4), 325–333. https://doi.org/10.1016/0167-6393(87)90007-0

Newman, R. S., & Ratner, N. B. (2007). The role of selected lexical factors on confrontation naming accuracy, speed, and fluency in adults who do and do not stutter. *Journal of Speech, Language, and Hearing Research, 50*(1), 196–213. DOI: 10.1044/1092-4388(2007/016)

Nikvarz, N., & Sabouri, S. (2022). Drug-induced stuttering: A comprehensive literature review. *World Journal of Psychiatry, 12*(2), 236–249. https://doi.org/10.5498/wjp.v12.i2.236

Ntourou, K., Conture, E. G., & Lipsey, M. W. (2011). Language abilities of children who stutter: A meta-analytical review. *American Journal of Speech-Language Pathology, 20*(3), 163–179. doi: 10.1044/1058-0360(2011/09-0102)

Ofoe, L. C., Anderson, J. D., & Ntourou, K. (2018). Short-term memory, inhibition, and attention in developmental stuttering: A meta-analysis. *Journal of Speech, Language, and Hearing Research, 61*(7), 1626–1648. DOI: 10.1044/2018_JSLHR-S-17-0372

Olander, L., Smith, A., & Zelaznik, H. (2010). Evidence that a motor timing deficit is a factor in the development of stuttering. *Journal of Speech, Language, and Hearing Research, 53*(4), 876–886. doi: 10.1044/1092-4388(2009/09-0007)

Oller, D. K., & Eilers, R. E. (1988). The role of audition in infant babbling. *Child Development, 59*(2), 441–449. https://doi.org/10.2307/1130323

Oomen, C. C., & Postma, A. (2002). Limitations in processing resources and speech monitoring. *Language and Cognitive Processes, 17*(2), 163–184. https://psycnet.apa.org/doi/10.1080/01690960143000010

Oyoun, H. A., El Dessouky, H., Shohdi, S., & Fawzy, A. (2010). Assessment of working memory in normal children and children who stutter. *Journal of American Science, 6*(11), 562–569.

Packman, A. (2012). Theory and therapy in stuttering: A complex relationship. *Journal of Fluency Disorders, 37*(4), 225–233. https://doi.org/10.1016/j.jfludis.2012.05.004

Packman, A., & Attanasio, J. S. (2017). *Theoretical issues in stuttering*. Routledge.

Packman, A., Code, C., & Onslow, M. (2007). On the cause of stuttering: Integrating theory with brain and behavioral research. *Journal of Neurolinguistics, 20*(5), 353–362. https://doi.org/10.1016/j.jneuroling.2006.11.001

Packman, A., Onslow, M., Coombes, T., & Goodwin, A. (2001). Stuttering and lexical retrieval. *Clinical Linguistics & Phonetics, 15*(6), 487–498. https://doi.org/10.1080/02699200110059114

Park, J., & Logan, K. J. (2015). The role of temporal speech cues in facilitating the fluency of adults who stutter. *Journal of Fluency Disorders, 46*, 41–55. https://doi.org/10.1016/j.jfludis.2015.07.001

Parrell, B., & Houde, J. (2019). Modeling the role of sensory feedback in speech motor control and learning. *Journal of Speech, Language, and Hearing Research, 62*(8S), 2963–2985. DOI: 10.1044/2019_JSLHR-S-CSMC7-18-0127

Parrell, B., Ramanarayanan, V., Nagarajan, S., & Houde, J. (2019). The FACTS model of speech motor control: Fusing state estimation and task-based control. *PLoS computational biology, 15*(9), e1007321. DOI: 10.1371/journal.pcbi.1007321

Patri, J. F., Diard, J., & Perrier, P. (2019). Modeling sensory preference in speech motor planning: A Bayesian modeling framework. *Frontiers in Psychology, 10*, Article 2339. https://doi.org/10.3389/fpsyg.2019.02339

Pavlovia. (n.d.). [Computer software]. Open Science Tools. https://pavlovia.org

Pawluk, E. J., & Koerner, N. (2013). A preliminary investigation of impulsivity in generalized anxiety disorder. *Personality and Individual Differences, 54*(6), 732–737. https://doi.org/10.1016/j.paid.2012.11.027

Peirce, J. W. (2009). Generating stimuli for neuroscience using PsychoPy. *Frontiers in Neuroinformatics, 2*, Article 10. https://doi.org/10.3389/neuro.11.010.2008

Pelczarski, K. M., & Yaruss, J. S. (2016). Phonological memory in young children who stutter. *Journal of Communication Disorders, 62*, 54–66. https://doi.org/10.1016/j.jcomdis.2016.05.006

Pellowski, M. W. (2011). Word-finding and vocabulary abilities of adults who do and do not stutter. *Contemporary Issues in Communication Science and Disorders, 38*(Fall), 126–134. https://doi.org/10.1044/cicsd_38_F_126

Pellowski, M., & Conture, E. (2005). Lexical priming in picture naming of young children who do and do not stutter. *Journal of Speech, Language, and Hearing Research, 48*(2), 278–294. https://doi.org/10.1044/1092-4388(2005/019)

Perez, H. R., & Stoeckle, J. H. (2016). Stuttering: Clinical and research update. *Canadian Family Physician, 62*(6), 479–484.

Perkins, W. H. (1990). What is stuttering? *Journal of Speech and Hearing Disorders, 55*(3), 370–382. https://doi.org/10.1044/jshd.5503.370

Perkins, W. H., Kent, R. D., & Curlee, R. F. (1991). A theory of neuropsycholinguistic function in stuttering. *Journal of Speech, Language, and Hearing Research, 34*(4), 734–752. https://doi.org/10.1044/jshr.3404.734

Peters, H. M., Hulstijn, W., & Van Lieshout, P. H. (2000). Recent developments in speech motor research into stuttering. *Folia Phoniatrica et Logopaedica, 52*(1–3), 103–119. DOI: 10.1159/000021518

Petrosino, L., Fucci, D., Gorman, P., & Harris, D. (1987). Midline and off-midline tongue and right- and left-hand vibrotactile thresholds of stutterers and normal-speaking individuals. *Perceptual and Motor Skills, 65*(1), 253–254. https://doi.org/10.2466/pms.1987.65.1.253

Port, R. (2003). Meter and speech. *Journal of Phonetics, 31*(3–4), 599–611. https://doi.org/10.1016/j.wocn.2003.08.001

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

Postma, A., & Kolk, H. (1992). Error monitoring in people who stutter: Evidence against auditory feedback defect theories. *Journal of Speech, Language, and Hearing Research, 35*(5), 1024–1032. https://doi.org/10.1044/jshr.3505.1024

Postma, A., & Kolk, H. (1993). The covert repair hypothesis. *Journal of Speech, Language, and Hearing Research, 36*(3), 472–487. https://doi.org/10.1044/jshr.3603.472

Postma, A., Kolk, H., & Povel, D. (1990). Speech planning and execution in stutterers. *Journal of Fluency Disorders, 15*(1), 49–59. https://doi.org/10.1016/0094-730X(90)90032-N

Prestes, R., Andrade, A. N. D., Santos, R. B. F., Marangoni, A. T., Schiefer, A. M., & Gil, D. (2017). Temporal processing and long-latency auditory evoked potential in stutterers. *Brazilian Journal of Otorhinolaryngology, 83*(2), 142–146. DOI: 10.1016/j.bjorl.2016.02.015

Putnam, A., & Ringel, R. (1976). A cineradiographic study of articulation in two talkers with temporarily induced oral sensory deprivation. *Journal of Speech and Hearing Research, 19*(2), 247–266. https://doi.org/10.1044/jshr.1902.247

R Core Team. (2022). *R: A language and environment for statistical computing* [Computer software]. R Foundation for Statistical Computing. https://www.R-project.org/

Ramus, F., Rosen, S., Dakin, S., Day, B., Castellote, J., White, S., & Frith, U. (2003). Theories of developmental dyslexia: Insights from a multiple case study of dyslexic adults. *Brain, 126*(4), 841–865. DOI: 10.1093/brain/awg076

Rapp, B., & Caramazza, A. (2002). Selective difficulties with spoken nouns and written verbs: A single case study. *Journal of Neurolinguistics, 15*(3–5), 373–402. https://doi.org/10.1016/S0911-6044(01)00040-9

Ratyńska, J., Szkiełkowska, A., Markowska, R., Kurkowski, M., Mularzuk, M., & Skarżyński, H. (2012). Immediate speech fluency improvement after application of the Digital

Speech Aid in stuttering patients. *Medical Science Monitor: International Medical Journal of Experimental and Clinical Research, 18*(1), CR9–CR14.

Regier, D. A., Kuhl, E. A., & Kupfer, D. J. (2013). The DSM-5: Classification and criteria changes. *World Psychiatry, 12*(2), 92–98. https://doi.org/10.1002/wps.20050

Reinisch, E., & Sjerps, M. J. (2013). The uptake of spectral and temporal cues in vowel perception is rapidly influenced by context. *Journal of Phonetics, 41*(2), 101–116. https://doi.org/10.1016/j.wocn.2013.01.002

Richels, C., Buhr, A., Conture, E., & Ntourou, K. (2010). Utterance complexity and stuttering on function words in preschool-age children who stutter. *Journal of Fluency Disorders, 35*(3), 314–331. https://doi.org/10.1016/j.jfludis.2010.06.001

Riley, G. (2009). *Stuttering Severity Instrument for Children and Adults* (4th ed.). Pro-Ed.

Riley, G., & Riley, J. (1980). Motoric and linguistic variables among children who stutter: A factor analysis. *Journal of Speech and Hearing Disorders, 45*(4), 504–513. https://doi.org/10.1044/jshd.4504.504

Riley, J., & Riley, J. (2000). A revised component model for diagnosing and treating children who stutter. *Contemporary Issues in Communication Science and Disorders, 27*, 188–199. https://doi.org/10.1044/cicsd_27_F_188

Ringel, R., & Steer, M. (1963). Some effects of tactile and auditory alterations on speech output. *Journal of Speech and Hearing Research, 6*, 369–378. https://doi.org/10.1044/jshr.0604.369

Robb, M., Lynn, W., & O'Beirne, G. (2013). An exploration of dichotic listening among adults who stutter. *Clinical Linguistics & Phonetics, 27*(9), 681–693. https://doi.org/10.3109/02699206.2013.808698

Roelofs, A. (2006). The influence of spelling on phonological encoding in word reading, object naming, and word generation. *Psychonomic Bulletin & Review, 13*(1), 33–37. https://doi.org/10.3758/BF03193809

Roelofs, A. P. A., & Ferreira, V. S. (2019). The architecture of speaking (pp. 35–50). In *The Oxford Handbook of Psycholinguistics* (2nd ed.). Cambridge, MA: MIT Press.

Rosen, S. (1992). Temporal information in speech: Acoustic, auditory and linguistic aspects. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences, 336*(1278), 367–373. https://doi.org/10.1098/rstb.1992.0070

Sahin, H., Krespi, Y., Yilmaz, A., & Coban, O. (2005). Stuttering due to ischemic stroke. *Behavioural Neurology, 16*(1), 37–39. https://doi.org/10.1155/2005/941926

Sares, A., Deroche, M., Shiller, D., & Gracco, V. (2019). Adults who stutter and metronome synchronization: Evidence for a nonspeech timing deficit. *Annals of the New York Academy of Sciences, 1449*(1), 56–69. DOI: 10.1111/nyas.14117

Sasisekaran, J. (2013). Nonword repetition and nonword reading abilities in adults who do and do not stutter. *Journal of Fluency Disorders, 38*(3), 275–289. DOI: 10.1016/j.jfludis.2013.06.001

Sasisekaran, J., & Byrd, C. (2013). Nonword repetition and phoneme elision skills in school-age children who do and do not stutter. *International Journal of Language & Communication Disorders, 48*(6), 625–639. doi: 10.1111/1460-6984.12035

Sasisekaran, J., & Weisberg, S. (2014). Practice and retention of nonwords in adults who stutter. *Journal of Fluency Disorders, 41*, 55–71. doi: 10.1016/j.jfludis.2014.02.004

Sasisekaran, J., Luc, F., Smyth, R., & Johnson, C. (2006). Phonological encoding in the silent speech of persons who stutter. *Journal of Fluency Disorders, 31*(1), 1–21. https://doi.org/10.1016/j.jfludis.2005.11.005

Schwartz, H., & Conture, E. (1988). Subgrouping young stutterers. *Journal of Speech, Language, and Hearing Research, 31*(1), 62–71. https://doi.org/10.1044/jshr.3101.62

Schwartz, M. F., & Dell, G. S. (2012). The dorsal stream contribution to phonological retrieval in object naming. *Brain and Language, 122*(3), 196–210. doi: 10.1093/brain/aws300

Schwartze, M., & Kotz, S. (2013). A dual-pathway neural architecture for specific temporal prediction. *Neuroscience & Biobehavioral Reviews, 37*(10), 2587–2596. https://doi.org/10.1016/j.neubiorev.2013.08.005

Schwartze, M., & Kotz, S. (2024). Timing patterns in the extended basal ganglia system. In H. Merchant & V. K. Balasubramaniam (Eds.), *Neurobiology of interval timing* (pp. 275–282). Springer.

Scott, C., & Ringel, R. (1971). The effects of motor and sensory disruptions on speech: A description of articulation. *Journal of Speech and Hearing Research, 14*(4), 819–828. https://doi.org/10.1044/jshr.1404.819

Seery, C. H., Watkins, R. V., Mangelsdorf, S. C., & Shigeto, A. (2007). Subtyping stuttering II: Contributions from language and temperament. *Journal of Fluency Disorders, 32*(3), 197–217. DOI: 10.1016/j.jfludis.2007.07.001

Shaiman, S., & Gracco, V. (2002). Task-specific sensorimotor interactions in speech production. *Experimental Brain Research, 146*(4), 411–418. DOI: 10.1007/s00221-002-1195-5

Shallice, T., & Evans, M. (1978). The involvement of the frontal lobes in cognitive estimation. *Cortex, 14*(2), 294–303. https://doi.org/10.1016/S0010-9452(78)80055-0

Shamma, S., Patel, P., Mukherjee, S., Marion, G., Khalighinejad, B., Han, C., ... Mesgarani, N. (2021). Learning speech production and perception through sensorimotor interactions. *Cerebral Cortex Communications, 2*(1), 1-18. https://doi.org/10.1093/texcom/tgaa091

Shannon, R. V., Zeng, F. G., Kamath, V., Wygonski, J., & Ekelid, M. (1995). Speech recognition with primarily temporal cues. *Science, 270*(5234), 303–304. https://doi.org/10.1126/science.270.5234.303

Shao, J., Bakhtiar, M., & Zhang, C. (2022). Impaired categorical perception of speech sounds under the backward masking condition in adults who stutter. *Journal of Speech, Language, and Hearing Research, 65*(7), 2554–2570. DOI: 10.1044/2022_JSLHR-21-00276

Shao, Z., Meyer, A., & Roelofs, A. (2013). Selective and nonselective inhibition of competitors in picture naming. *Memory & Cognition, 41*(8), 1200–1211. https://doi.org/10.3758/s13421-013-0332-7

Shao, Z., Roelofs, A., & Meyer, A. (2012). Sources of individual differences in the speed of naming objects and actions: The contribution of executive control. *Quarterly Journal of Experimental Psychology, 65*(10), 1927–1944. DOI: 10.1080/17470218.2012.670252

Shao, Z., Roelofs, A., Acheson, D., & Meyer, A. (2014). Electrophysiological evidence that inhibition supports lexical selection in picture naming. *Brain Research, 1586*, 130–142. https://doi.org/10.1016/j.brainres.2014.07.009

SheikhBahaei, S., Millwater, M., & Maguire, G. (2023). Stuttering as a spectrum disorder: A hypothesis. *Current Research in Neurobiology, 5*, 100116. https://doi.org/10.1016/j.crneur.2023.100116

Shelton, J., & Caramazza, A. (1999). Deficits in lexical and semantic processing: Implications for models of normal language. *Psychonomic Bulletin & Review, 6*(1), 5–27. DOI: 10.3758/bf03210809

Shetty, H. (2016). Temporal cues and the effect of their enhancement on speech perception in older adults – A scoping review. *Journal of Otology, 11*(3), 95–101. doi: 10.1016/j.joto.2016.08.001

Shinn, J. (2003). Temporal processing: The basics. *The Hearing Journal, 56*(7), 52. DOI: 10.1097/01.HJ.0000292557.52409.67

Sikorskii, I. A. (1889). O zaikanii [On stuttering]. St. Petersburg: Izdanie Karla Rikkera.

Siritzky, E. M., Cox, P. H., Nadler, S. M., Grady, J. N., Kravitz, D. J., & Mitroff, S. R. (2023). Standard experimental paradigm designs and data exclusion practices in cognitive psychology can inadvertently introduce systematic "shadow" biases in participant samples. *Cognitive Research: Principles and Implications, 8*(1), 66. DOI: 10.1186/s41235-023-00520-y

Slis, A., Savariaux, C., Perrier, P., & Garnier, M. (2023). Rhythmic tapping difficulties in adults who stutter: A deficit in beat perception, motor execution, or sensorimotor integration? *PloS ONE, 18*(2), e0281194. DOI: 10.1371/journal.pone.0276691

Smith, A. (1990). Toward a comprehensive theory of stuttering: A commentary. *Journal of Speech and Hearing Disorders, 55*(3), 398–401. https://doi.org/10.1044/jshd.5503.398

Smith, A. (1999). Stuttering: A unified approach to a multifactorial, dynamic disorder. In N. B. Ratner & C. E. Healey (Eds.), *Stuttering research and practice: Bridging the gap* (pp. 27–44). Psychology Press.

Smith, A., & Kleinow, J. (2000). Kinematic correlates of speaking rate changes in stuttering and normally fluent adults. *Journal of Speech, Language, and Hearing Research, 43*(2), 521–536. https://doi.org/10.1044/jslhr.4302.521

Smith, A., Goffman, L., Sasisekaran, J., & Weber-Fox, C. (2012). Language and motor abilities of preschool children who stutter: Evidence from behavioral and kinematic indices of

nonword repetition performance. *Journal of Fluency Disorders, 37*(4), 344–358. https://doi.org/10.1016/j.jfludis.2012.06.001

Smith, A., Sadagopan, N., Walsh, B., & Weber-Fox, C. (2010). Increasing phonological complexity reveals heightened instability in inter-articulatory coordination in adults who stutter. *Journal of Fluency Disorders, 35*(1), 1–18. doi: 10.1016/j.jfludis.2009.12.001

Smith, K. A., Iverach, L., O'Brian, S., Kefalianos, E., & Reilly, S. (2014). Anxiety of children and adolescents who stutter: A review. *Journal of Fluency Disorders, 40*, 22–34. https://doi.org/10.1016/j.jfludis.2014.01.003

Smits-Bandstra, S., De Nil, L., & Rochon, E. (2006). The transition to increased automaticity during finger sequence learning in adult males who stutter. *Journal of Fluency Disorders, 31*(1), 22–42. https://doi.org/10.1016/j.jfludis.2005.11.004

Snodgrass, J., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory, 6*(2), 174–215. https://doi.org/10.1037/0278-7393.6.2.174

Spencer, C., & Weber-Fox, C. (2014). Preschool speech articulation and nonword repetition abilities may help predict eventual recovery or persistence of stuttering. *Journal of Fluency Disorders, 41*, 32–46. https://doi.org/10.1016/j.jfludis.2014.06.001

St. Onge, K., & Calvert, J. (1964). Stuttering research. *Quarterly Journal of Speech, 50*(2), 159–165. https://doi.org/10.1080/00335636409382657

Stewart, C., Evans, W., & Fitch, J. (1985). Oral form perception skills of stuttering and nonstuttering children measured by stereognosis. *Journal of Fluency Disorders, 10*(1), 25–42. https://doi.org/10.1016/0094-730X(85)90029-4

Sugathan, N., & Maruthy, S. (2020). Nonword repetition and identification skills in Kannada speaking school-aged children who do and do not stutter. *Journal of Fluency Disorders, 63*, Article 105745. https://doi.org/10.1016/j.jfludis.2019.105745

Sutton, S., & Chase, R. (1961). White noise and stuttering. *Journal of Speech and Hearing Research, 4*(1), 72–75. https://doi.org/10.1044/jshr.0401.72

Tani, T., & Sakai, Y. (2011). Analysis of five cases with neurogenic stuttering following brain injury in the basal ganglia. *Journal of Fluency Disorders, 36*(1), 1–16. https://doi.org/10.1016/j.jfludis.2010.12.002

Temple, C. M. (1997). Cognitive neuropsychology and its application to children. *Journal of Child Psychology and Psychiatry, 38*(1), 27–52. DOI: 10.1111/j.1469-7610.1997.tb01504.x

Therapeutes.com. (n.d.). Bégaiement : cause et traitements naturels efficaces. https://www.therapeutes.com/affections/troubles-de-la-voix/begaiement

Theys, C., Jaakkola, E., Melzer, T. R., De Nil, L. F., Guenther, F. H., Cohen, A. L., ... Joutsa, J. (2024). Localization of stuttering based on causal brain lesions. *Brain, 147*(6), 2203–2215. https://doi.org/10.1093/brain/awae059

Thomas, J. G., & Sharp, P. (2019). Mechanistic science: A new approach to comprehensive psychopathology research that relates psychological and biological phenomena. *Clinical Psychological Science, 7*(2), 196–215. https://psycnet.apa.org/doi/10.1177/2167702618810223

Thomas, M. (2003). Multiple causality in developmental disorders: Methodological implications from computational modelling. *Developmental Science, 6*(5), 537–556. https://psycnet.apa.org/doi/10.1111/1467-7687.00311

Thomas, M., & Karmiloff-Smith, A. (2002). Are developmental disorders like cases of adult brain damage? Implications from connectionist modelling. *Behavioral and Brain Sciences, 25*(6), 727–750. DOI: 10.1017/s0140525x02000134

Thomas, M., & Karmiloff-Smith, A. (2005). Can developmental disorders reveal the component parts of the human language faculty? *Language Learning and Development, 1*(1), 65–92. https://doi.org/10.1207/s15473341lld0101_5

Thompson, C. (2006). Single subject controlled experiments in aphasia: The science and the state of the science. *Journal of Communication Disorders, 39*(4), 266–291. https://doi.org/10.1016/j.jcomdis.2006.02.003

Tichenor, S., Leslie, P., Shaiman, S., & Yaruss, J. S. (2018). Speaker and observer perceptions of physical tension during stuttering. *Folia Phoniatrica et Logopaedica, 69*(4), 180–189. DOI: 10.1159/000486032

Tilsen, S. (2013). A dynamical model of hierarchical selection and coordination in speech planning. *PloS ONE, 8*(4), e62800. https://doi.org/10.1371/journal.pone.0062800

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

Tourville, J., Reilly, K., & Guenther, F. (2008). Neural mechanisms underlying auditory feedback control of speech. *NeuroImage, 39*(3), 1429–1443. https://doi.org/10.1016/j.neuroimage.2007.09.054

Towgood, K., Meuwese, J., Gilbert, S., Turner, M. S., & Burgess, P. (2009). Advantages of the multiple case series approach to the study of cognitive deficits in autism spectrum disorder. *Neuropsychologia, 47*(13), 2981–2988. https://doi.org/10.1016/j.neuropsychologia.2009.06.028

Townsend, J., & Ashby, F. (1978). Methods of modeling capacity in simple processing systems. In J. Castellan & F. Restle (Eds.), *Cognitive theory* (Vol. 3, pp. 200–239). Erlbaum.

Townsend, J., & Ashby, F. (1983). *Stochastic modeling of elementary psychological processes*. Cambridge University Press.

Träff, U., Olsson, L., Östergren, R., & Skagerlund, K. (2017). Heterogeneity of developmental dyscalculia: Cases with different deficit profiles. *Frontiers in Psychology, 7*, Article 1910. doi: 10.3389/fpsyg.2016.02000

Tremblay, S., Shiller, D., & Ostry, D. (2003). Somatosensory basis of speech production. *Nature, 423*(6942), 866–869. https://doi.org/10.1038/nature01710

Tsiamtsiouris, J., & Cairns, H. (2009). Effects of syntactic complexity and sentence-structure priming on speech initiation time in adults who stutter. *Journal of Fluency Disorders, 34*(2), 108–123. DOI: 10.1044/1092-4388(2009/08-0063)

Vahab, M., Shojaei, K., Ahmadi, A., & Nasiri, M. (2014). Phonological working memory in 4-8 year-old Persian children who stutter. *Journal of Fluency Disorders, 42*, 20–26. https://doi.org/10.30476/jrsr.2014.41062

van de Vorst, R., & Gracco, V. (2017). Atypical non-verbal sensorimotor synchronization in adults who stutter may be modulated by auditory feedback. *Journal of Fluency Disorders, 53*, 14–25. https://doi.org/10.1016/j.jfludis.2017.05.004

van Lieshout, P. (2017). Coupling dynamics in speech gestures: Amplitude and rate influences. *Experimental Brain Research, 235*(8), 2495–2510. DOI: 10.1007/s00221-017-4983-7

Van Lieshout, P., Starkweather, C., Hulstijn, W., & Peters, H. (1995). Effects of linguistic correlates of stuttering on EMG activity in nonstuttering speakers. *Journal of Speech, Language, and Hearing Research, 38*(2), 360–372. https://doi.org/10.1044/jshr.3802.360

van Lieshout, P., Hulstijn, W., & Peters, H. (1996). From planning to articulation in speech production: What differentiates a person who stutters from a person who does not stutter? *Journal of Speech, Language, and Hearing Research, 39*(3), 546–564. https://doi.org/10.1044/jshr.3903.546

van Lieshout, P., Ben-David, B., Lipski, M., & Namasivayam, A. (2014). The impact of threat and cognitive stress on speech motor control in people who stutter. *Journal of Fluency Disorders, 40*, 93–109. https://doi.org/10.1016/j.jfludis.2014.02.003

Van Riper, C. (1971). *The nature of stuttering* (1st ed.). Prentice-Hall.

Van Riper, C. (1982). *The nature of stuttering* (2nd ed.). Prentice-Hall.

Vasic, N., & Wijnen, F. (2005). Stuttering as a monitoring deficit. In R. J. Hartsuiker, R. Bastiaanse, A. Postma, & F. Wijnen (Eds.), *Phonological encoding and monitoring in normal and pathological speech* (pp. 238–259). Psychology Press.

Van Thorre, A., Patri, J. F., & Vannuscorps, G. (2025). Testing theories of stuttering: Cognitive heterogeneity but shared orosensory tactile weaknesses. *Cognitive Neuropsychology*, 1–40. https://doi.org/10.1080/02643294.2025

Wagner, M., & Watson, D. (2010). Experimental and theoretical advances in prosody: A review. *Language and Cognitive Processes, 25*(7), 905–945. doi: 10.1080/01690961003589492

Weber-Fox, C. (2001). Neural systems for sentence processing stuttering. *Journal of Speech, Language, and Hearing Research, 44*(4), 814–825. https://doi.org/10.1044/1092-4388(2001/064)

Weber-Fox, C., Spencer, R., Spruill, J., & Smith, A. (2004). Phonologic processing in adults who stutter. *Journal of Speech, Language, and Hearing Research, 47*(6), 1244–1258. DOI: 10.1044/1092-4388(2004/094)

Webster, R., & Lubker, B. (1968). Interrelationships among fluency producing variables in stuttered speech. *Journal of Speech and Hearing Research, 11*(4), 754–766. https://doi.org/10.1044/jshr.1104.754

Webster, W. (1990). Evidence in bimanual finger-tapping of an attentional component to stuttering. *Behavioural Brain Research, 37*(2), 93–100. https://doi.org/10.1016/0166-4328(90)90084-r

Westphal, G. (1933). An experimental study of certain motor abilities of stutterers. *Child Development, 4*(3), 214–227. https://doi.org/10.2307/1125683

Wheeldon, L. R., & Levelt, W. J. (1995). Monitoring the time course of phonological encoding. *Journal of Memory and Language, 34*(3), 311–334. https://doi.org/10.1006/jmla.1995.1014

White, P. (1990). Ideas about causation in philosophy and psychology. *Psychological Bulletin, 108*(1), 3–18. https://doi.org/10.1037/0033-2909.108.1.3

White, S., Milne, E., Rosen, S., Hansen, P., Swettenham, J., Frith, U., & Ramus, F. (2006). The role of sensorimotor impairments in dyslexia: A multiple case study of dyslexic children. *Developmental Science, 9*(3), 237–255. https://doi.org/10.1111/j.1467-7687.2006.00483.x

Wieland, E., McAuley, J., Dilley, L., & Chang, S. (2015). Evidence for a rhythm perception deficit in children who stutter. *Brain and Language, 144*, 26–34. doi: 10.1016/j.bandl.2015.03.008

Williams, B. (2010). Perils of evidence-based medicine. *Perspectives in Biology and Medicine, 53*(1), 106–120. https://doi.org/10.1353/pbm.0.0132

Wiltshire, C. E., Chesters, J., Krishnan, S., Healy, M. P., & Watkins, K. E. (2020). An fMRI study of initiation and inhibition of manual responses in people who stutter. *bioRxiv*, 2020–09. https://doi.org/10.1162/IMAG.a.89

Wiltshire, C., Cler, G., Chiew, M., Freudenberger, J., Chesters, J., Healy, M. P., Hoole, P., & Watkins, K. (2024). Speaking to a metronome reduces kinematic variability in typical speakers and people who stutter. *PLOS ONE, 19*(10), e0309612. https://doi.org/10.1371/journal.pone.0309612

Wingate, M. (1970). Effect on stuttering of changes in audition. *Journal of Speech and Hearing Research, 13*(4), 861–873. https://doi.org/10.1044/jshr.1304.861

Wingate, M. (1970). Stuttering, 1970: Where do we stand? *Journal of Speech and Hearing Research, 13*(1), 5–8. DOI: 10.1044/jshr.1301.05

Wingate, M. (1988). *The structure of stuttering: A psycholinguistic analysis*. Springer.

Wingate, M. (1964). A standard definition of stuttering. *Journal of Speech and Hearing Disorders, 29*(4), 484–489. https://doi.org/10.1044/jshd.2904.484

Wingate, M. (1988). The Fault Line. In *The structure of stuttering: A psycholinguistic analysis* (pp. 179–185). Springer.

Wohl, M. (1968). The electronic metronome: An evaluative study. *International Journal of Language & Communication Disorders, 3*(1), 89–98.

Woodward, J. (2005). *Making things happen: A theory of causal explanation*. Oxford University Press.

Woumans, E., Van der Linden, L., Hartsuiker, R., Duyck, W., Moerenhout, C., de Partz, M. P., ... Szmalec, A. (2021). Speech fluency in bilinguals who stutter: Language proficiency and attentional demands as mediating factors. *Journal of Fluency Disorders, 69*, Article 105850. https://doi.org/10.1016/j.jfludis.2021.105850

Yairi, E. (1990). Subtyping child stutterers for research purposes. *ASHA Reports Series*, (18), 40–53.

Yairi, E. (2007). Subtyping stuttering I: A review. *Journal of Fluency Disorders, 32*(3), 165–196. https://doi.org/10.1016/j.jfludis.2007.04.001

Yairi, E., & Ambrose, N. (2013). Epidemiology of stuttering: 21st century advances. *Journal of Fluency Disorders, 38*(2), 66–87. https://doi.org/10.1016/j.jfludis.2012.11.002

Yairi, E., & Seery, C. H. (2021). *Stuttering: Foundations and clinical applications* (3rd ed.). Plural Publishing.

Yaruss, J. S. (1997). Clinical measurement of stuttering behaviors. *Contemporary Issues in Communication Science and Disorders, 24*(Spring), 27–38.

Yaruss, J., LaSalle, L., & Conture, E. (1998). Evaluating stuttering in young children. *American Journal of Speech-Language Pathology, 7*(4), 62–76. https://doi.org/10.1044/1058-0360.0704.62

Zalewska, A. (2014, March). Collaboration at CERN: Greater than the sum of its parts. *CERN Courier*. https://cerncourier.com/a/collaboration-at-cern-greater-than-the-sum-of-its-parts/

Zelaznik, H., Smith, A., & Franz, E. (1994). Motor performance of stutterers and nonstutterers on timing and force control tasks. *Journal of Motor Behavior, 26*(4), 340–347. https://doi.org/10.1080/00222895.1994.9941690

Zelaznik, H., Smith, A., Franz, E., & Ho, M. (1997). Differences in bimanual coordination associated with stuttering. *Acta Psychologica, 96*(3), 229–243. DOI: 10.1016/s0001-6918(97)00014-0

Zhao, L., & Lian, M. (2021). Lexical planning in people who stutter: A defect in lexical encoding or the planning scope? *Frontiers in Psychology, 12*, Article 581304. https://doi.org/10.3389/fpsyg.2021.581304

Zorowitz, S., & Niv, Y. (2023). Improving the reliability of cognitive task measures: A narrative review. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging, 8*(8), 789–797. DOI: 10.1016/j.bpsc.2023.02.004

Zsilavecz, U. (1981). Cybernetic functioning in stuttering. *South African Journal of Communication Disorders, 28*(1), 60–67. https://doi.org/10.4102/sajcd.v28i1.353

Appendix

The following appendix contains supplementary information related to Chapter 4, along with additional data supporting the consistency of orosensory tactile difficulties exhibited by some of the twelve AWS assessed in the study described in that chapter.

1. Supplementary Information

Supplementary Information for

Testing theories of stuttering: cognitive heterogeneity but shared orosensory tactile weaknesses

Amélie Van Thorre, Jean-François Patri & Gilles Vannuscorps
Correspondence should be addressed to Gilles Vannuscorps, Email: gilles.vannuscorps@uclouvain.be

This document includes:

Supplementary Table S1

Supplementary Figures S1-S9

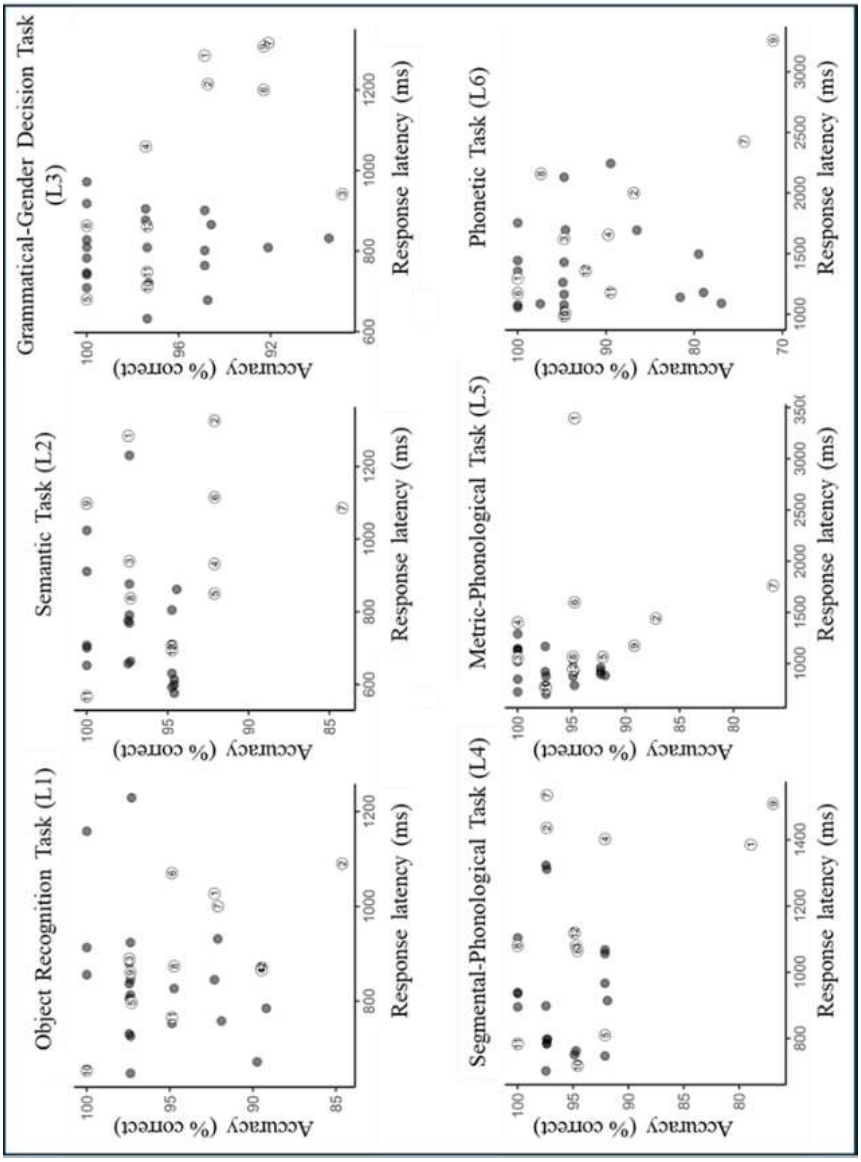
Supplementary Table S1

History of therapy and strategies

Participants	Period of intervention	Strategies used at the time of testing
AWS 1	Early adulthood	Slowing speech and frequent pausing
AWS 2	Early adulthood	Slowing speech and vowel prolongation
AWS 3	Childhood	None reported
AWS 4	Childhood and early adulthood	Slowing speech, use of fillers
AWS 5	Early adulthood	Slowing speech, frequent pausing, and avoidance of triggers
AWS 6	Early adulthood	Slowing speech and frequent pausing
AWS 7	Childhood and early adulthood	Slowing speech and frequent pausing
AWS 8	Beginning of adolescence	Use of fillers, word substitution
AWS 9	Early adulthood	Slowing speech
AWS 10	End of childhood	None reported
AWS 11	Early adulthood	Slowing speech
AWS 12	Early adulthood	Slowing speech and frequent pausing

Supplementary Figure S1

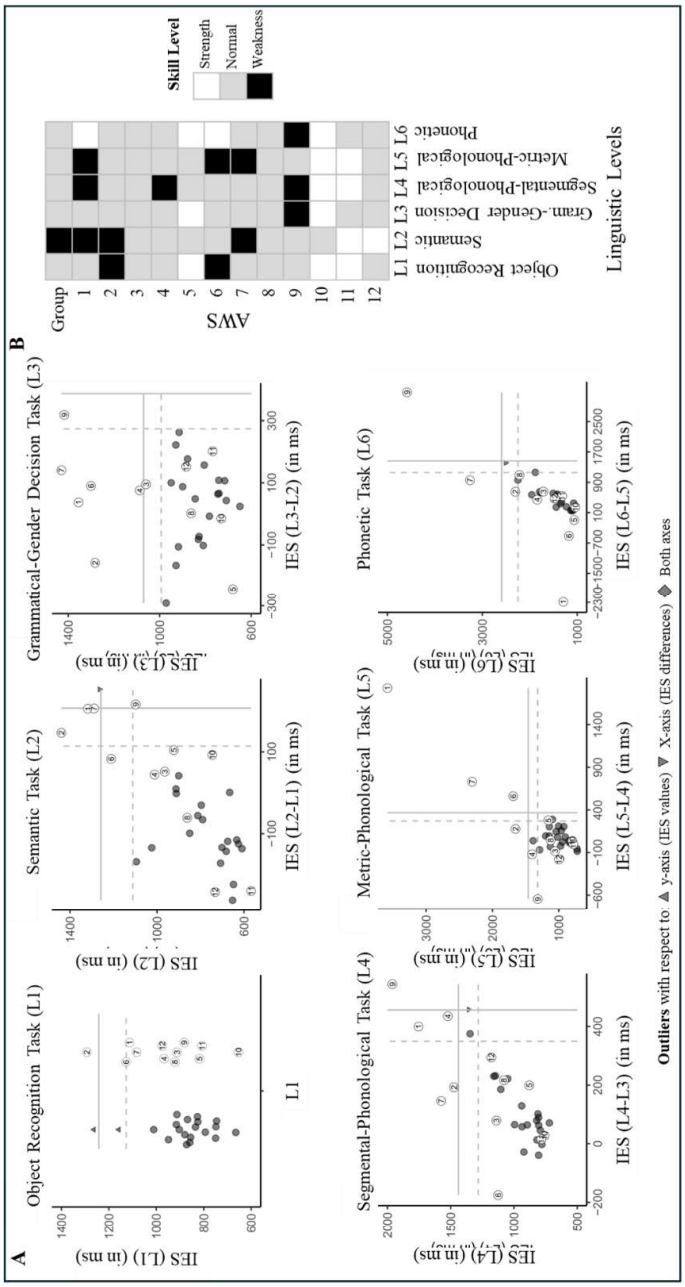
Linguistic assessment: accuracy and response latencies



Note. Scatter plots of accuracy (y-axis) and response time (x-axis) for the six linguistic tasks are presented for controls (gray-filled circles) and AWS (white circles labeled with their ID).

Supplementary Figure S2

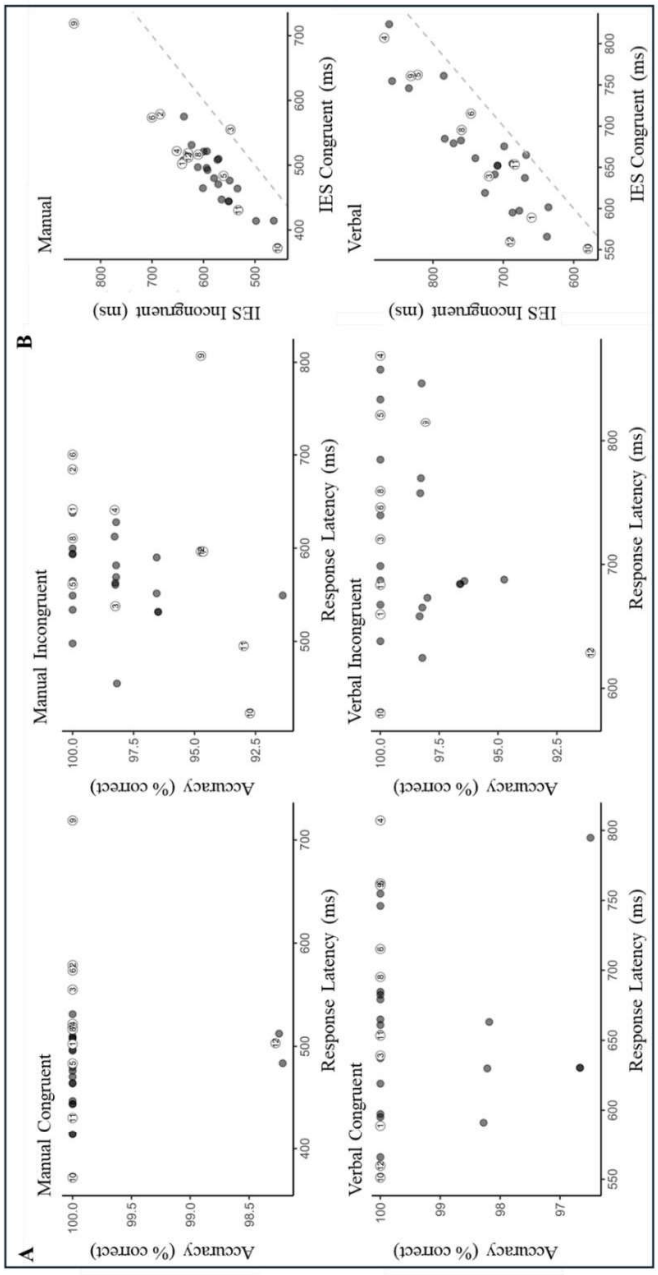
Results of the linguistic assessments analyzed without excluding outliers



Note. A. Scatter plots of performances for the six tasks of the linguistic assessments are shown for controls (grey filled data, triangles indicating outliers) and AWS (white circles labeled with their ID). Grey lines (vertical and horizontal) indicate statistical thresholds, corresponding to the first (plain; $p = 0.01$) and fifth (dashed; $p = 0.05$) estimated percentiles of the control distribution for each axis. B. Overall performance of each participant in all the tasks administered obtained from the combined performances in efficiency and transitioning cost. As detailed in the text, performances of AWS exceeding the fifth percentile threshold along both axes (i.e. falling within the upper right quadrant defined by the dashed lines) are considered as a weakness, performances of AWS numerically below the controls' median along both axes indicate a strength. In all other cases, the performance is classified as normal.

Supplementary Figure S3

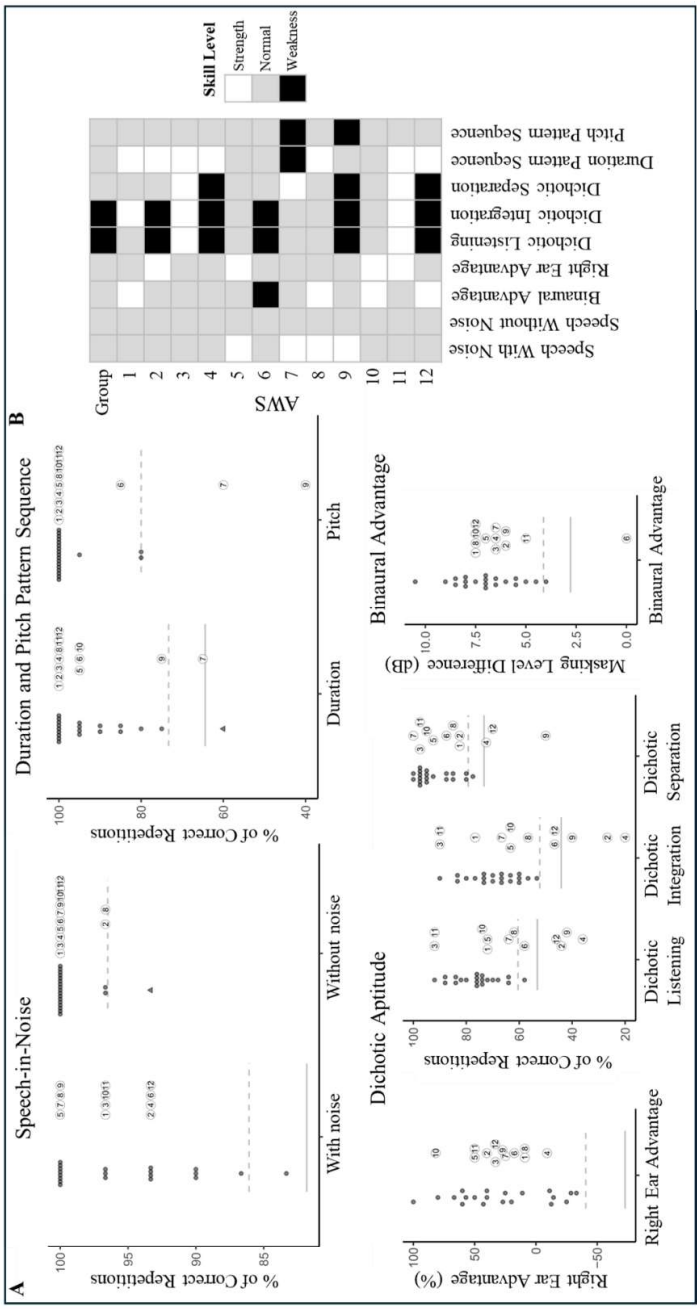
Results of inhibitory control assessment: accuracy, response latencies and efficiency by condition



Note. A. Scatter plots of accuracy (y-axis) and response time (x-axis) for the two modalities (manual and verbal) and the two conditions (congruent and incongruent) of the Flanker task are presented for controls (gray-filled circles) and AWS (white circles labeled with their ID). B. Participants' inverse efficiency scores (IES), with incongruent trial scores on the y-axis and congruent trial scores on the x-axis. Higher values indicate lower efficiency in processing, with points deviating from the diagonal suggesting differential effects of congruency on task performance.

Supplementary Figure S4

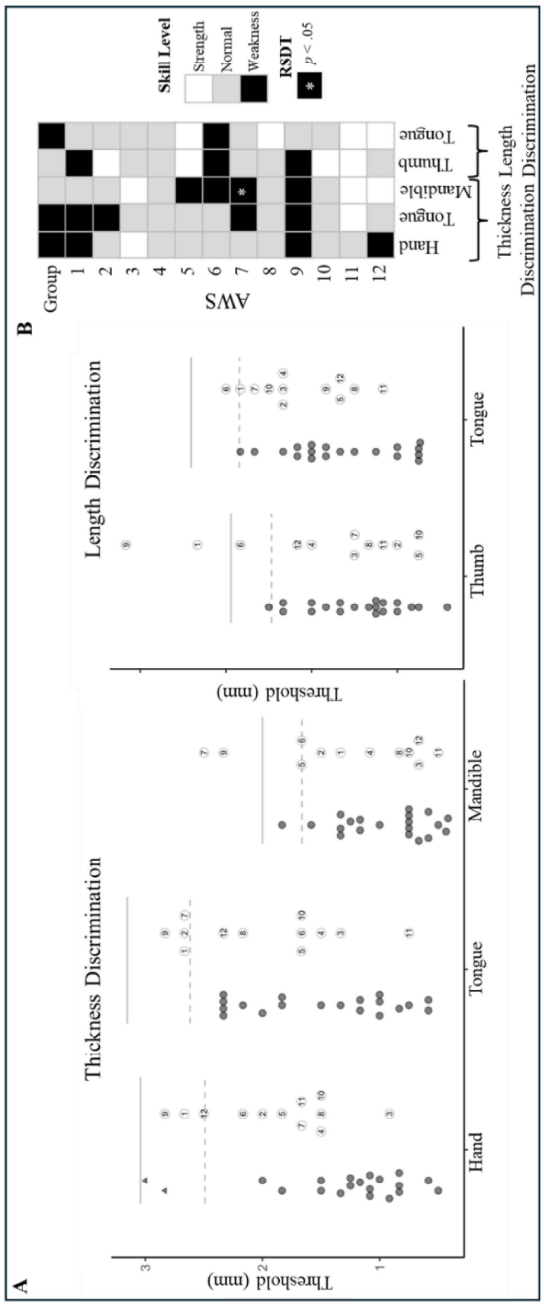
Results of auditory assessments analyzed without excluding outliers



Note. A. Performance of controls (grey filled circles, triangles indicating outliers) and AWS (white circles labeled with their ID) in the auditory assessment. Grey lines indicate statistical thresholds, corresponding to the first (plain) and fifth (dashed) estimated percentiles of the control distribution. B. Summary of the results of the AWS. Performances below the fifth percentile threshold are considered as a weakness, performances of AWS numerically above the controls' median indicate a strength. In all other cases, the performance is classified as normal.

Supplementary Figure S5

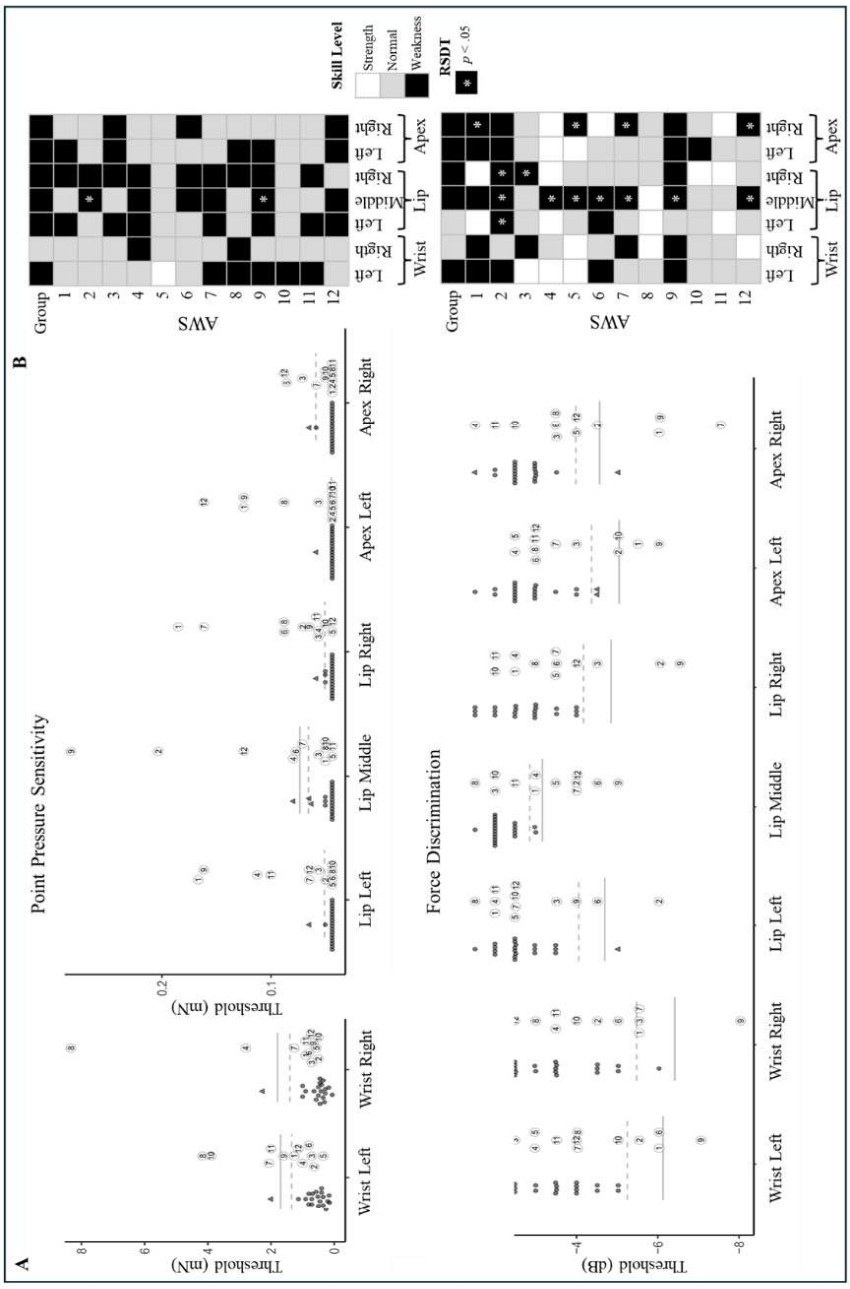
Results of the proprioceptive assessments analyzed without excluding outliers



Note. A. Performance of controls (grey filled circles, triangles indicating outliers) and AWS (white circles labeled with their ID) in proprioceptive assessments. Grey lines indicate statistical thresholds, corresponding to the first (plain) and fifth (dashed) estimated percentiles of the control distribution. B. Summary of the results. Performances below the fifth percentile threshold are considered as a weakness, performances of AWS numerically above the controls' median indicate a strength. In all other cases, the performance is classified as normal. Cases displaying weaknesses in speech related orofacial structures (black cells for Tongue or the Mandible) were compared to their respective performances in control body parts (Hand and Thumb). The * symbol indicates when a significant difference ($p < .05$) was obtained in at least one of the corresponding RSDT tests (Crawford & Garthwaite, 2005).

Supplementary Figure S6

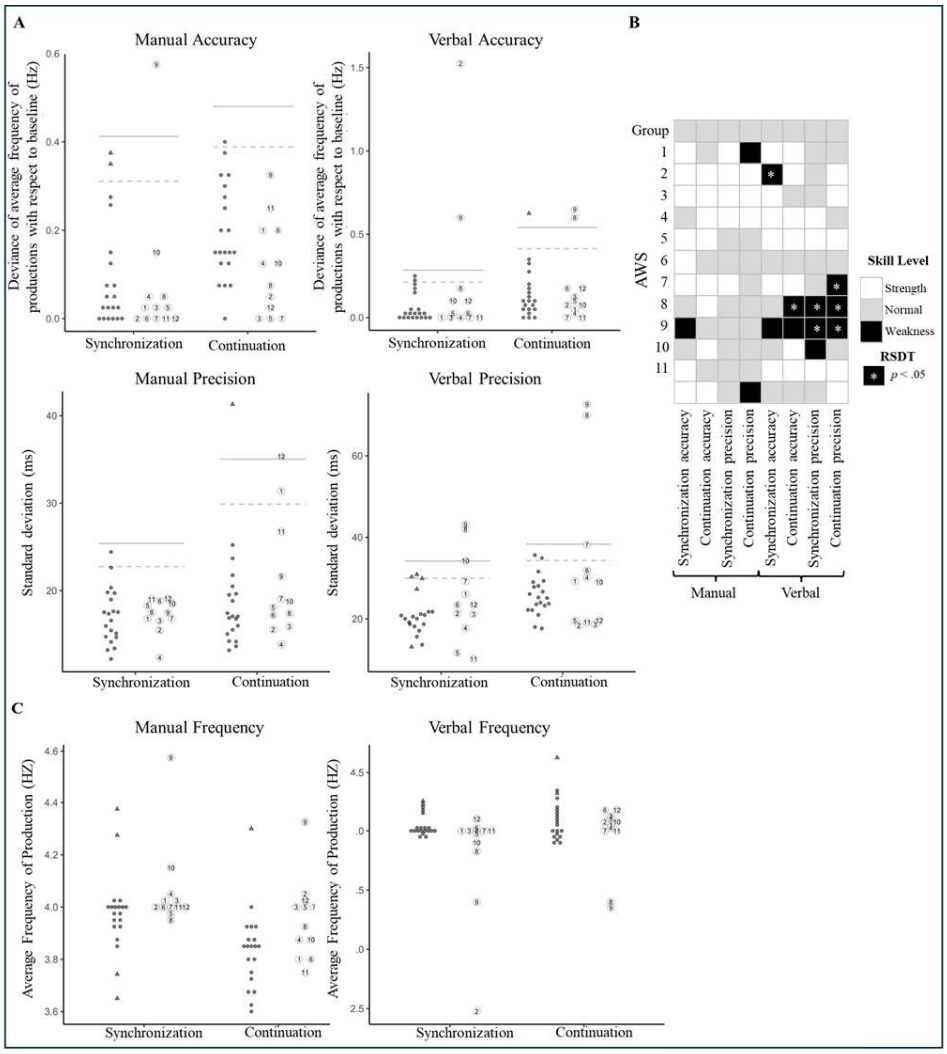
Results of the tactile assessments analyzed without excluding outliers



Note. A. Performance of controls (grey filled circles, triangles indicating outliers) and AWS (white circles labeled with their ID) in the tactile assessments. Grey lines indicate statistical thresholds, corresponding to the first (plain) and fifth (dashed) estimated percentiles of the control distribution. B. Summary of the results. Performances below the fifth percentile threshold are considered as a weakness, performances of AWS numerically above the controls' median indicate a strength. In all other cases, the performance is classified as normal. Cases displaying a weakness in speech related orofacial structures (black cells in Lip and Apex regions) were compared to their respective performances in control body parts (Left and Right Wrists). The * symbol indicates when a significant difference ($p < .05$) was obtained in at least one of the two corresponding RSDT tests (Crawford & Garthwaite, 2005). Note that for Point Pressure Sensitivity, RSDT comparisons were not conducted, as the control group's interquartile ranges for the Lip and Apex regions were zero.

Supplementary Figure S7

Results of the rhythmic assessments analyzed without excluding outliers

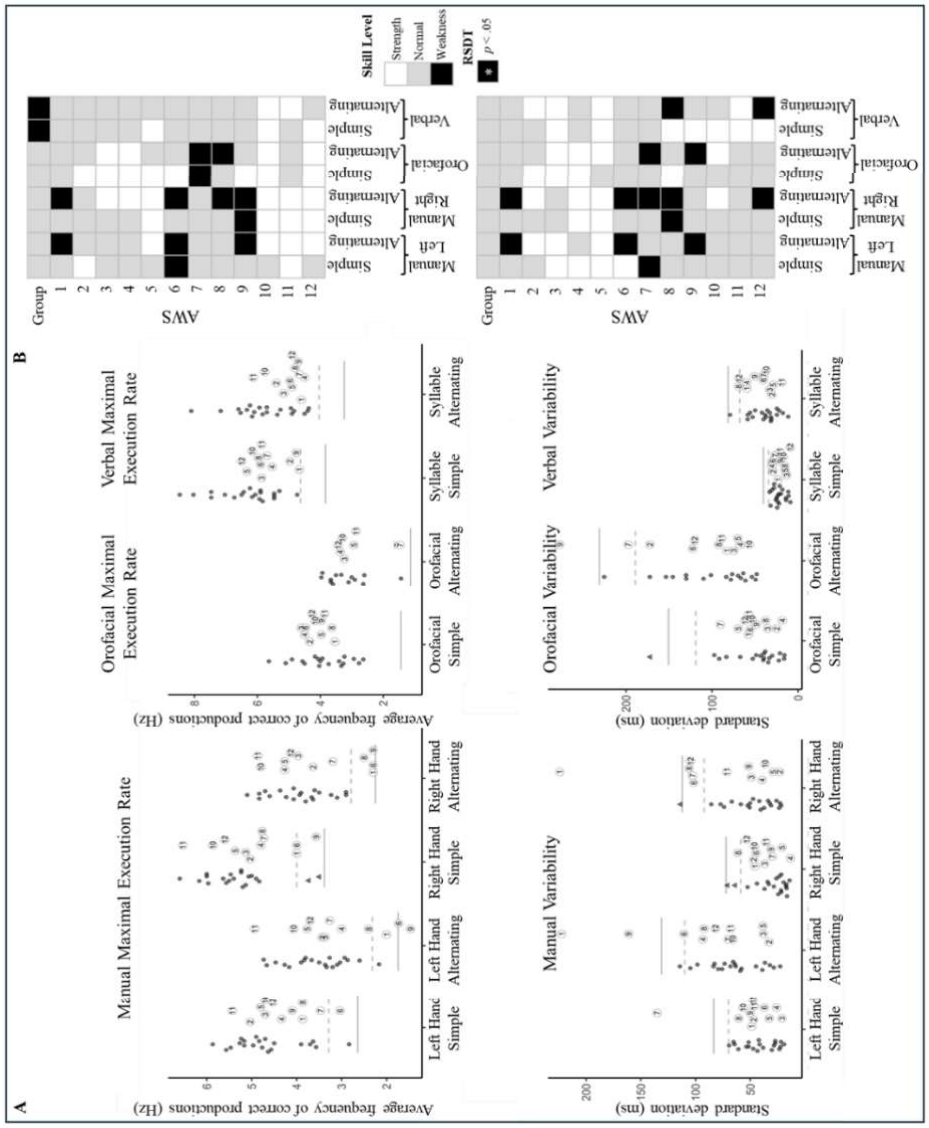


Note. A. Performances of controls (grey filled circles, triangles indicating outliers) and AWS (white circles labeled with their ID) are shown for the two measures of rhythmic abilities, accuracy of the frequency production from the expected 4 Hz baseline (upper panels) and precision of production (lower panels), for finger movement (left panels) and speech-related movement (right panels) tasks. Grey lines indicate statistical thresholds, corresponding to the

first (plain) and fifth (dashed) estimated percentiles of the control distribution. B. Summary of the results. Performances below the fifth percentile threshold are considered as a weakness, performances of AWS numerically above the controls' median indicate a strength. In all other cases, the performance is classified as normal. Cases displaying difficulties in speech related zones (dark cells in Verbal modality) were compared to their respective performances in the associated control condition (Manual modality). The * symbol indicates when a significant difference ($p < .05$) was obtained in the corresponding RSDT test (Crawford & Garthwaite, 2005). C. Average frequency of production for finger movement (left panel) and speech-related movement (right panel).

Supplementary Figure S8

Results of the motor assessments analyzed without excluding outliers

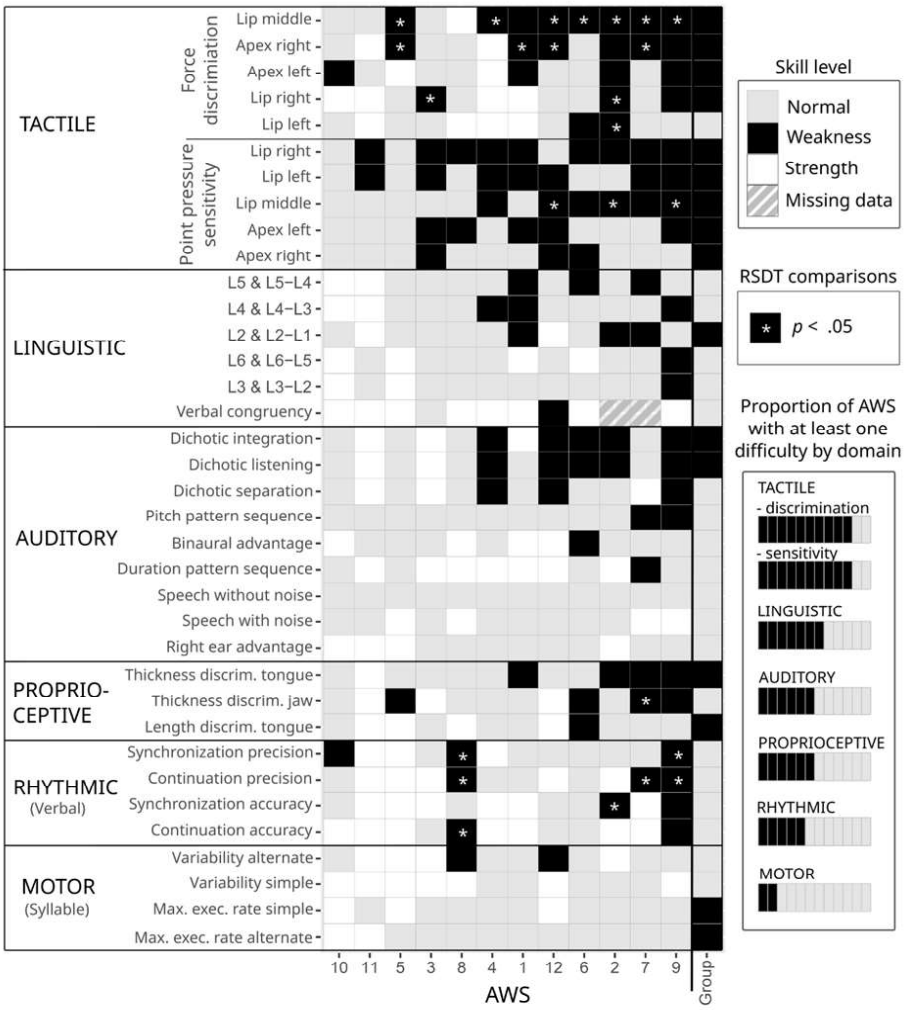


Note. A. Performances of controls (grey filled circles, triangles indicating outliers) and AWS (white circles labeled with their ID) are shown for the two measures of motor abilities, maximal execution rate (left panels) and variability of production (right panels), for the

manual (upper panel), the verbal and the orofacial modalities (lower panel). Grey lines indicate statistical thresholds, corresponding to the first (plain) and fifth (dashed) estimated percentiles of the control distribution. B. Summary of the results. Performances below the fifth percentile threshold are considered as a weakness, performances of AWS numerically above the controls' median indicate a strength. In all other cases, the performance is classified as normal. Cases displaying difficulties in speech related movements (Verbal production in simple and alternating conditions) were compared to their respective performances in control modalities (Orofacial and Manual movements). No significant differences were found in any of the RSDT comparisons (all *ps* > .05).

Supplementary Figure S9

Summary of the results in all speech related assessments without excluding outliers



Note. Performances of each AWS and overall group (horizontal axis) in all speech related assessments (vertical axis) are summarized into three categories: "weakness" (performances below the 5th percentile; black cells), "strength" (performances above the 50th percentile) and "normal" (performances between the 5th and 50th percentile). Domains are sorted from top to bottom according to the proportion of weaknesses observed across AWS. Subjects are sorted from left to right according to the proportion of weaknesses obtained across speech related

tasks. RSDT refers to the Revised Standardized Difference test (Crawford & Garthwaite, 2005).

2. Performance consistency

To evaluate the consistency of our orofacial tactile empirical findings, seven AWS (1, 3, 6, 9, 10, 11, and 12) previously tested in Chapter 4 were retested by our team on measures of tactile detection and force discrimination, utilizing a different protocol.

In the experimental design presented in Chapter 4, we attempted to measure the threshold at which each participant "changed response from 'no sensation' to 'yes, I perceive something'" (Lawless & Heymann, 2010, p.125) using a method of limits. However, limitations of this method have been identified. There is considerable variability in participants' response strategies: some may set high criteria and require strong certainty before responding, while others may report any minimal sensation. This variability introduces inconsistencies and raises concerns about whether the method of limits accurately measures tactile thresholds or is confounded by individual differences in response bias. Consequently, we cannot be entirely certain that our results reflect true tactile thresholds rather than subjective differences in reporting criteria (see Lawless & Heyman, 2010). To address this issue, researchers often implement a forced-choice paradigm. Accordingly, in the new protocol, we utilized a variant of the method of constant stimuli, presenting stimuli at fixed intensity levels independently of participants' previous answers and assessing their detection probabilities.

Point pressure sensitivity and force discrimination were reassessed, focusing on two speech-relevant orofacial structures—the right and left lower lip and the right and left apex of the tongue—as well as on one non-orofacial site, the fingertip. The selection of tested body parts was partially revised to address challenges encountered during the initial assessment. In the first testing session, we examined the right and left corners of the mouth. However, this area proved particularly difficult to assess, especially in male participants with facial hair, as it was challenging to avoid applying the filament to body hair rather than the skin. For this reason, we chose to test the right and left lower lip in the new design. In the initial assessment, the wrist was selected as the control zone. This choice was motivated by observations of reduced fingertip sensitivity in some participants, likely attributable to

activities such as manual labor or playing musical instruments. However, during testing, substantial variation in sensitivity was also observed at the wrists, particularly among participants who habitually wore watches or bracelets. As both sites could be influenced by participants' daily habits, we ultimately chose the fingertip as the control zone. This decision was further supported by literature indicating that fingertip mechanoreceptive properties more closely approximate those of orofacial regions, rendering it a more appropriate control zone for studies of orofacial tactile sensitivity (Haggard & Boer, 2014).

Participants' probability of detecting von Frey filaments (Fruhstorfer, Gross, & Selbmann, 2001), calibrated to produce a contact force of 0.088 mN for orofacial sites and 0.063 mN for the fingertip, was assessed to evaluate **point pressure sensitivity**. These specific forces were selected because they are typically perceived by most individuals, thereby minimizing the likelihood that participants' responses were due to chance or response bias. Consequently, if a participant who stutters was unable to reliably detect the stimulation, this would indicate genuine difficulties in tactile perception.

The experiment employed a method of constant stimuli with a three-interval forced-choice paradigm and was controlled using PsychoPy software (version 2023.3) (Peirce, 2009). Testing was divided into three blocks of 35 trials each (7 trials for each body part), resulting in a total of 105 trials. Participants were instructed to keep their eyes closed for the duration of the task. Each trial was initiated by a sequence of three brief tones (200 ms each), spaced 100 ms apart, signaling the beginning of the trial. Subsequently, three intervals (each lasting 1500 ms) were indicated by a brief tone at the onset of each interval. During one of these intervals, the experimenter delivered a tactile stimulus. At the end of the three intervals, a brief high-pitched tone signaled the conclusion of the trial and prompted participants to indicate in which interval the stimulation had occurred by pressing one of three buttons on a computer mouse: the left button for the first interval, the central button for the second, and the right button for the third interval. This forced-choice design required participants to select a response in every trial, thereby minimizing potential response biases, for example, avoiding the tendency of more conservative participants to withhold responses when uncertain. The use of three intervals (rather than two) set the probability of correct responses due to chance at 33% (1 in 3), offering a stricter criterion for performance than a two-interval design (which has a 50% chance level). This lower chance level makes the task more sensitive for detecting

genuine differences in tactile detection performance, ensuring that above-chance performance more likely reflects true perceptual ability rather than guessing. During the first block, a familiarization phase was conducted prior to testing each body part. This phase consisted of four trials for the fingertip and the apex of the tongue, and six trials for the lip.

For each body part, the detection probability of each AWS was calculated and compared to the mean performance of the control group (n = 40).

Participants' **force discrimination** was assessed using von Frey filaments (Fruhstorfer, Gross, & Selbmann, 2001) calibrated to deliver contact forces of 0.5 mN and 2 mN for the tongue, and 0.71 mN and 2 mN for the fingertip and the lip. This procedure enabled us to evaluate participants' ability to discriminate between two slightly different contact pressures. These specific forces were chosen because they are sufficiently intense to be perceived by all participants, thereby minimizing the likelihood that responses were attributable to an inability to detect one of the filaments. Consequently, if a participant who stutters was unable to reliably discriminate between the two stimuli, this would indicate a genuine difficulty in force discrimination.

The experiment employed a method of constant stimuli and was controlled using PsychoPy software (version 2023.3) (Peirce, 2009). Testing was divided into three blocks of 35 trials each (7 trials for each body part), resulting in a total of 105 trials. Participants were instructed to keep their eyes closed for the duration of the task. Each trial began with a sequence of three brief tones (200 ms each), spaced 100 ms apart, signaling the start of the trial. The onset and the end of each of the two intervals (each lasting 1500 ms) was indicated by a brief tone. During the first interval, the experimenter delivered one of the two possible tactile stimulations—either with the weaker or the stronger filament. During the second interval, the other stimulation was applied. At the end of the two intervals, participants were prompted to indicate in which interval the stimulation was the strongest by pressing one of two buttons on a computer mouse: the left button for the first interval and the right button for the second interval. In the first block, a familiarization phase of four trials was administered before testing each body part.

For each body part, the detection probability of each AWS was calculated and compared to the mean performance of the control group (n = 34).

The main results of the point pressure sensitivity and force discrimination tasks utilizing the new protocol are presented in Figures 1. To assess AWS performance, we employed Crawford and Howell's modified t-test (1998), which yields a p-value that estimates the proportion of control participants with performances below the tested score (Crawford & Garthwaite, 2006). With the exception of AWS 11, the remaining six AWS who were retested exhibited difficulties in one or more orofacial sites in at least one of the two tasks. For five out of six AWS, some of these orofacial difficulties were consistent with those identified in the assessment conducted in Chapter 4 (see Table 2).

Furthermore, we employed Crawford's Bayesian Standardized Difference Test (BSDT; Crawford & Garthwaite, 2007) to investigate dissociations between performance on orofacial sites and a control body part (the fingertip). The BSDT assesses whether a difference in performance between two conditions in an individual AWS exceeds what would be expected based on the control group's performance in these same conditions. Notably, for 5 out of 7 AWS (AWS 1, 6, 9, 10, 12), there was at least one statistically significant difference between their performance on an orofacial structure (tongue or lip) and the control site (fingertip), with all BSDT p-values < .033.