



**HAL**  
open science

# Interaction entre structure rythmique et sens d'agentivite en production de la parole

Jinyu Li

► **To cite this version:**

Jinyu Li. Interaction entre structure rythmique et sens d'agentivite en production de la parole. Linguistique. Université de la Sorbonne nouvelle - Paris III, 2023. Français. NNT : 2023PA030119 . tel-04611770

**HAL Id: tel-04611770**

**<https://theses.hal.science/tel-04611770>**

Submitted on 14 Jun 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Sorbonne  
Nouvelle  
université des cultures



**UNIVERSITÉ SORBONNE NOUVELLE - CNRS**  
École Doctorale 622 : SCIENCES DU LANGAGE  
UMR7018 / Laboratoire de Phonétique et Phonologie

Thèse de doctorat en  
**PHONÉTIQUE, PHONOLOGIE ET SCIENCE DE LA PAROLE**

# **INTERACTION ENTRE STRUCTURE RYTHMIQUE ET SENS D'AGENTIVITÉ EN PRODUCTION DE LA PAROLE**

Présentée par  
**Jinyu LI**

Sous la direction de  
**Cécile FOUGERON**  
**Leonardo LANCIA**

Soutenue à Paris le 19 décembre 2023 devant le jury composé de :

Rapporteur	<b>Pascal PERRIER</b>	PR, Grenoble INP/Université Grenoble Alpes
Rapportrice	<b>Corine ASTESANO</b>	PR, Université Toulouse - Jean Jaurès
Examineur	<b>Louis GOLDSTEIN</b>	PR, University of Southern California
Examineur	<b>Noël NGUYEN</b>	PR, Aix-Marseille Université
Examinatrice	<b>Yao YAO</b>	Pr.Assist., Hong Kong Polytechnic University
Directrice	<b>Cécile FOUGERON</b>	DR, CNRS/Université Sorbonne Nouvelle
Directeur	<b>Leonardo LANCIA</b>	CR, CNRS/Aix-Marseille Université



---

# INTERACTION ENTRE STRUCTURE RYTHMIQUE ET SENS D'AGENTIVITÉ EN PRODUCTION DE LA PAROLE

Pour s'adapter aux imprévus lors de la production de la parole, le système moteur intègre les informations sensorielles (ex. le retour auditif), et bénéficie du groupement rythmique, qui est caractérisé par la prosodie. Cependant, le système sensorimoteur d'un locuteur traite différemment les événements acoustiques relatifs à sa propre voix par rapport à ceux d'autrui. Cette thèse vise à examiner la flexibilité de la production de la parole en analysant le rôle organisateur de la prosodie et la sensation subjective d'un locuteur sur le contrôle de sa voix (c.-à-d. le sens d'agentivité relatif à sa voix).

Des expériences des perturbations du retour auditif ont été menées chez des locutrices francophones. Avec le retour auditif retardé (DAF), la différence de durée entre les voyelles accentuées et non accentuées s'est accrue, soulignant une plus grande flexibilité durant la production des accents. De plus, le DAF a induit une réorganisation du rythme de la parole avec un renforcement du groupement syllabique. Avec un décalage continu de la fréquence fondamentale ( $f_0$ ) du retour auditif, la majorité des locutrices ont aligné leur  $f_0$  sur celle modifiée du retour auditif, suggérant que leur système sensorimoteur aurait traité la voix perçue comme une entrée externe. La présence simultanée du DAF et d'un décalage de la  $f_0$  a entraîné une réduction des effets du DAF par rapport à la condition sans décalage de la  $f_0$ . Cette observation suggère une réduction du sens d'agentivité relatif à la voix chez les locutrices, ainsi qu'une interaction entre l'organisation rythmique et le sens d'agentivité dans les processus sensorimoteurs de production de la parole.

**Mots-Clés :** contrôle sensorimoteur de la parole, retour auditif, rythme de la parole, structure prosodique, sens d'agentivité, alignement phonétique, traitement de la parole

---

# INTERACTION BETWEEN RHYTHMIC STRUCTURE AND SENSE OF AGENCY IN SPEECH PRODUCTION

To adapt to unforeseen circumstances during speech production, the motor system integrates sensory information (e.g., auditory feedback) and benefits from rhythmic grouping, which is characterized by prosody. However, a speaker's sensorimotor system processes acoustic events related to their own voice differently from those of others. This thesis aims to examine the flexibility of speech production by analyzing the organizing role of both prosody and a speaker's subjective sensation of control over his voice (i.e., the sense of agency related to his voice).

Experiments of auditory feedback perturbations were conducted with French-speaking female speakers. With delayed auditory feedback (DAF), the duration difference between accented and unaccented vowels increased, highlighting greater flexibility during accent production. Furthermore, DAF induced a reorganization of speech rhythm with enhanced syllabic grouping. With a constant shift in the fundamental frequency ( $f_0$ ) of auditory feedback, the majority of female speakers aligned their  $f_0$  with the modified auditory feedback, suggesting that their sensorimotor system processed the perceived voice as an external input. The simultaneous presence of DAF and an  $f_0$  shift resulted in a reduction of DAF effects compared to the condition without an  $f_0$  shift. This observation suggests a reduction in the sense of agency related to the voice among female speakers, as well as an interaction between rhythmic organization and sense of agency in sensorimotor processes of speech production.

**Key words:** sensorimotor control of speech, auditory feedback, speech rhythm, prosodic structure, sense of agency, phonetic alignment, speech processing

---

## REMERCIEMENTS

Je tiens à exprimer ma profonde reconnaissance à Leonardo Lancia et Cécile Fougeron pour leurs orientations, leurs conseils, leurs commentaires et leur confiance tout au long de ce travail. Je souhaite tout particulièrement remercier Leonardo, qui m'a constamment encouragé à réfléchir, à m'exprimer et à échanger, et qui a toujours fait preuve d'une grande patience pour m'enseigner de nouvelles connaissances et méthodes, répondre à toutes mes questions et m'aider à résoudre des problèmes.

Je suis très honorée que Pascal Perrier, Corine Astesano, Louis Goldstein, Noël Nguyen et Yao Yao fassent partie du jury de cette thèse.

### Un grand merci,

à Rachid et Takeki pour leurs précieux commentaires et conseils lors de chaque comité de suivi.

à Cécile et Leonardo (encore une fois) de m'avoir offert l'opportunité de travailler sur leurs projets, qui non seulement ont permis de financer ce travail, mais m'ont également donné l'occasion de participer à divers échanges scientifiques.

à tous les commentaires précieux que j'ai reçus lors des conférences.

à Clémence, Amélie, Alice, Chloé, Alexis, Maddie et Madame Christelle Porte pour leur relecture d'une partie de cette thèse.

à Gwenaëlle et Roland pour leur disponibilité et leur aide tout au long de ces années.

à toutes les participantes de mes expériences.

à tous les collègues du LPP (Angéline, Daria, Amélie, Clémence, Alexis, Maddie, Andrés, Alice ...) qui ont égayé nos déjeuners et pauses goûter en partageant des blagues et anecdotes.

à Minmin, Dayeon, Boram et Lila, qui sont non seulement mes collègues, mais aussi de véritables amies. Avec elles, j'ai partagé des moments de joie et de tristesse, où nous avons su nous soutenir mutuellement. Leur présence est toujours source d'encouragement pour moi.

à docteur Hao et futur docteur Jarvis pour leur écoute attentive, leurs partages précieux d'expérience et leur encouragement constant tout au long de la réalisation de ce travail.

à Chloé, Angèle, Yinxin, Xueqing, qui ont enrichi mon expérience à Paris et en France.

à mes amis Jin, Biyun, Xiaoting, Xi, Yishu, Ruoyu, pour leur sollicitude à distance au fil des années.

à mon grand-père et ma grand-mère. Je regrette profondément de ne pas avoir pu être à leurs côtés avant qu'ils ne s'en aillent vers un autre monde.

à mes parents, dont la compréhension de mon travail est limitée, mais qui font preuve d'un soutien indéfectible.

à la ville de Paris, qui, bien qu'elle ne soit pas parfaite, m'offre constamment la possibilité de découvrir la beauté de son architecture, l'art, la musique, la gastronomie ...

---

à moi-même pour avoir surmonté les hésitations, le stress, les déceptions et la fatigue, tout en conservant une attitude positive.

à la vie.

---

# TABLE DES MATIÈRES

<b>GLOSSAIRE.....</b>	<b>XIII</b>
<b>INTRODUCTION GENERALE.....</b>	<b>1</b>
1. Contexte .....	1
2. Structure de la thèse .....	3
<b>CHAPITRE I : COMPLEXITE DE LA PRODUCTION DE LA PAROLE.....</b>	<b>5</b>
<b>1. Groupement rythmique réduisant la complexité du contrôle moteur .....</b>	<b>6</b>
1.1 Articulateurs travaillant en synergie .....	6
1.2 Émergence du rythme .....	7
1.3 Rythme syllabique.....	9
1.4 Hiérarchisation du groupement rythmique .....	10
<b>2. Lien et différences entre la prosodie et le rythme .....</b>	<b>12</b>
2.1 Accentuation, métrique et unités prosodiques.....	13
2.2 Unités prosodiques du français .....	14
2.2.1 Application de l'accentuation en français .....	14
2.2.2 Système métrique du français.....	15
2.2.3 Caractéristiques de l'accentuation du français .....	16
2.3 Formalisation des phénomènes métriques.....	17
2.4 Rythme dérivant de la prosodie tout en s'en distinguant.....	18
2.5 Mesurer des régularités rythmiques .....	19
<b>3. Prosodie en tant que mécanisme stabilisant du contrôle moteur.....</b>	<b>21</b>
3.1 Rythme en tant que système coordinatif.....	21
3.2 Coordination entre les composantes rythmiques à différentes échelles temporelles .....	22
3.2.1 Composante syllabique et composante supra-syllabique .....	22
3.2.2 Coordination entre la composante syllabique et la composante supra-syllabique .....	23
<b>4. Résumé du chapitre .....</b>	<b>25</b>
<b>CHAPITRE II : FLEXIBILITE DE LA PRODUCTION DE LA PAROLE .....</b>	<b>29</b>
<b>1. Contrôle sensorimoteur pendant la production de la parole .....</b>	<b>30</b>
1.1 Processus du contrôle sensorimoteur dans la production de la parole .....	30
1.1.1 Intégration sensorimotrice .....	30
1.1.2 Intégration des retours sensoriels et du mécanisme prédictif.....	31
1.2 Retour auditif suivi pendant la production de la parole.....	33
1.2.1 Dimensions du retour auditif suivies.....	33
1.2.2 Intégration du retour auditif et somatosensoriel .....	35
<b>2. Contrôle sensorimoteur de la production de la parole interagissant avec la complexité linguistique .....</b>	<b>36</b>
2.1 Contrôle sensorimoteur modulé par l'expérience linguistique.....	36
2.2 Rôle de l'accentuation dans le contrôle sensorimoteur.....	37
2.3 Interaction entre organisation rythmique et contrôle sensorimoteur en production de la parole : perspectives.....	38
<b>3. Rôle du sens d'agentivité dans le contrôle sensorimoteur pendant la production de la parole .....</b>	<b>39</b>
3.1 Effet des entrées externes sur la production de la parole .....	39
3.1.1 Convergence phonétique .....	39
3.1.2 Cibles sensorimotrices .....	40
3.2 Génération du sens d'agentivité implicite .....	41
3.2.1 Définition du sens d'agentivité.....	41
3.2.2 Lien entre le sens d'agentivité implicite et le contrôle sensorimoteur.....	42
3.2.3 Génération du sens d'agentivité implicite pendant la production de la parole.....	44

3.3 Flexibilité du sens d'agentivité implicite .....	48
3.3.1 Compétition entre les cibles externes et internes .....	48
3.3.2 Altération du sens d'agentivité implicite.....	49
3.4 Sens d'agentivité implicite modulant le contrôle sensorimoteur de la parole .....	50
<b>4. Résumé du chapitre .....</b>	<b>51</b>
<b>CHAPITRE III : MESURER LES PATRONS RYTHMIQUES DE LA PAROLE PRODUITE SOUS PERTURBATION(S) AUDITIVES(S).....</b>	<b>53</b>
<b>1. Questions de recherche et hypothèses.....</b>	<b>54</b>
1.1 Question de recherche I et hypothèses .....	54
1.2 Question de recherche II et hypothèses.....	57
<b>2. Expérimentation.....</b>	<b>59</b>
2.1 Expérience I Retour auditif retardé à différents niveaux (expérience pilote).....	59
2.1.1 Objectives.....	59
2.1.2 Participants.....	60
2.1.3 Matériaux.....	60
2.1.4 Procédure.....	61
2.1.5 Traitement des données.....	62
2.2 Expérience II Partie A – condition uniquement DAF .....	64
2.2.1 Objectives.....	64
2.2.2 Participants.....	64
2.2.3 Matériaux.....	64
2.2.4 Procédure.....	65
2.2.5 Traitement des données.....	66
2.3 Expérience II Partie B – condition perturbations multiples .....	66
2.3.1 Objectives.....	66
2.3.2 Participants, matériaux, procédure et traitement des données.....	66
<b>3. Méthodes d'analyses .....</b>	<b>67</b>
3.1 Mesurer la différence d'allongement entre les syllabes accentuées et les syllabes non accentuées sous l'effet du DAF.....	68
3.2 Mesurer la coordination entre la composante syllabique et la composante supra-syllabique ..	71
3.2.1 Présentation générale de la méthode.....	71
3.2.2 Extraction de la composante syllabique et la composante supra-syllabique .....	72
3.2.3 Normalisation d'amplitude .....	73
3.2.4 Calcul de l'indice de coordination – Phase Locking Value (PLV).....	75
<b>4. Résumé du chapitre .....</b>	<b>76</b>
<b>CHAPITRE IV : ANALYSE DU ROLE DE L'ACCENTUATION DANS LE CONTROLE SENSORIMOTEUR PENDANT LA PRODUCTION DE LA PAROLE.....</b>	<b>77</b>
<b>1. Comparer l'effet d'allongement du DAF sur la durée des voyelles dans différentes positions .....</b>	<b>78</b>
1.1 Résultats de l'expérience I .....	78
1.2 Résultats de l'expérience II (« condition uniquement DAF »).....	80
1.3 Conclusion.....	82
<b>2. Analyse du rôle organisateur de l'accentuation .....</b>	<b>83</b>
2.1 Effet du DAF sur la différence de durée entre les voyelles accentuables et non accentuables – expérience I.....	83
2.2 Effet du DAF sur la différence de durée entre les voyelles accentuées et non accentuées – expérience II (« condition uniquement DAF »).....	85
2.3 Effets d'allongement du DAF interagissant avec l'allongement final .....	87
<b>3. Discussion du chapitre .....</b>	<b>88</b>
<b>CHAPITRE V : ANALYSE DU ROLE DU RYTHME DANS LA STABILISATION DU CONTROLE SENSORIMOTEUR PENDANT LA PRODUCTION DE LA PAROLE .....</b>	<b>91</b>
<b>1. Effet du DAF sur la coordination des composantes rythmiques dans les phrases de différents niveaux de difficulté de production .....</b>	<b>92</b>

2. Effet du DAF sur la coordination des composantes rythmiques dans les phrases avec différente complexité syllabique .....	94
2.1 Coordination des composantes rythmiques au niveau de la phrase .....	94
2.2 Coordination des composantes rythmiques au niveau du syntagme accentuel .....	96
2.3 Observations complémentaires : variation entre les locutrices .....	98
3. PLV et groupement syllabique .....	101
3.1 Nombre de pics syllabiques et supra-syllabiques .....	101
3.2 Effet du DAF lié au nombre de pics .....	103
4. Discussion du chapitre .....	105
4.1 Établissement de la stabilité de coordination des composantes rythmiques .....	105
4.2 Coordination des composantes rythmiques dans différents syntagmes accentuels .....	107
4.3 Variabilité individuelle .....	108
4.4 Implications des résultats de l'effet du DAF sur l'organisation rythmique .....	109
4.5 Discussion sur la méthode de PLV .....	110
<b>CHAPITRE VI : ÉVALUATION DU SENS D'AGENTIVITE ET SON ROLE ORGANISATEUR DANS LE CONTROLE SENSORIMOTEUR .....</b>	<b>113</b>
1. Différence individuelle en réponse au décalage continu de la $f_0$ du retour auditif .....	114
2. Relation entre alignement phonétique et sens d'agentivité .....	119
2.1 Comparer la différence de durée entre la syllabe accentuée et non accentuée dans chaque syntagme accentuel .....	119
2.1.1 Comparaison entre la condition avec DécalF0 et la condition sans DécalF0 .....	120
2.1.2 Comparaison entre les locutrices dans la « condition perturbations multiples » et celles dans la « condition uniquement DAF » .....	121
2.1.3 Conclusion de la section .....	124
2.2 Comparer le degré de coordination entre les composantes rythmiques .....	124
2.2.1 Comparaison entre la condition avec DécalF0 et la condition sans DécalF0 .....	125
2.2.2 Comparaison entre les locutrices dans la « condition perturbations multiples » et celles dans la « condition uniquement DAF » .....	128
2.2.3 Conclusion de la section .....	130
2.3 Comparer les effets indirects du DAF .....	131
2.3.1 Comparaison de l'effet du DAF sur la $f_0$ des voyelles .....	131
2.3.2 Comparaison de l'effet du DAF sur l'intensité des voyelles .....	132
2.3.3 Conclusion de la section .....	135
3. Discussion du chapitre .....	135
3.1 Alignement phonétique en situation d'altération du retour auditif et d'interaction verbale .....	135
3.2 Réduction du sens d'agentivité sous le décalage continu de la $f_0$ du retour auditif .....	137
3.2.1 Mesurer le sens d'agentivité .....	137
3.2.2 Réduction du sens d'agentivité et interaction entre le sens d'agentivité et la prosodie .....	137
3.2.3 Sens d'agentivité en continuum .....	138
<b>CHAPITRE VII : CONCLUSION GENERALE ET PERSPECTIVES .....</b>	<b>141</b>
1. Conclusion .....	141
2. Perspectives .....	148
2.1 Rôle de l'accentuation dans le contrôle sensorimoteur pendant la production de la parole ..	148
2.2 Organisation rythmique dans le contrôle sensorimoteur pendant la production de la parole .....	149
2.3 Flexibilité du sens d'agentivité pendant la production de la parole .....	150
2.4 Modulation de la flexibilité du contrôle sensorimoteur dans la production de la parole .....	150
<b>REFERENCES .....</b>	<b>153</b>
<b>ANNEXE .....</b>	<b>165</b>

---

## TABLE DES FIGURES

Figure 1 Illustration de l'alignement des régularités rythmiques entre les articulateurs, ainsi qu'entre le domaine articulatoire et le domaine acoustique, adaptée de Poeppel et Assaneo (2020) .....	8
Figure 2 Illustration de la relation entre la production de la parole et le rythme de la parole .....	9
Figure 3 Illustration des niveaux du groupement rythmique .....	12
Figure 4 Illustration de la grille métrique étendue extrait de Di Cristo (2016) .....	18
Figure 5 Illustration de la corrélation entre le groupement syllabique et la relation temporelle entre la composante syllabique et la composante supra-syllabique .....	25
Figure 6 Schéma illustrant une modalité possible de l'intégration d'un mécanisme prédictif et des retours sensoriels dans le système de production de la parole (adapté de Parrell et al., 2019) .....	32
Figure 7 Schéma représentant le lien entre le sens d'agentivité et le contrôle sensorimoteur .....	43
Figure 8 Illustration du lien entre l'activité du cortex auditif et la génération du sens d'agentivité, ainsi que nos hypothèses concernant le rôle du sens d'agentivité dans la régulation de la production de la parole .....	47
Figure 9 Illustration des procédures de l'expérience I .....	62
Figure 10 Recensement du nombre de répétitions contenant des erreurs de production dans chaque type de phrase à différents niveaux du DAF dans l'expérience I .....	63
Figure 11 Exemple du signal acoustique et la courbe f0 d'un essai dans l'expérience II .....	70
Figure 12 Modulations d'amplitude pour la phrase « Vivien vit le vin » (/vivjẽ vi lə vẽ/) .....	73
Figure 13 Procédure de normalisation .....	74
Figure 14 Durée des voyelles dans différentes positions, respectivement pour la phrase 1 (/valmõ vule vvar lə vøl/) et la phrase 3 (/mamã e mami mani ma mẽ/), dans l'expérience I .....	78
Figure 15 Durée des voyelles dans différentes positions, respectivement pour chaque type de phrase (CV, CVC, CCV(C)), dans l'expérience II (« condition uniquement DAF ») .....	81
Figure 16 Moyenne et écart-type du ratio entre la durée totale des voyelles accentuables et non accentuables pour la phrase 1 (/valmõ vule vvar lə vøl/) et la phrase 3 (/mamã e mami mani ma mẽ/) dans chaque position des essais expérimentaux dans l'expérience I .....	83
Figure 17 Moyenne et écart-type du ratio entre la durée de la voyelle accentuée et celle de la voyelle non-accentuée dans chaque SA (V2/V1 et V5/V4) pour les locutrices dans la « condition uniquement DAF » dans l'expérience II, respectivement pour chaque catégorie d'analyse (cf. chapitre III 3.1) .....	86
Figure 18 Durée des voyelles accentuées dans différentes positions des phrases .....	88
Figure 19 Durée des voyelles V4 et V5 dans les phrases non finales (catégorie d'analyse 1) à divers niveaux DAF (distingués par la couleur) .....	88

---

Figure 20 Moyenne et écart-type de la PLV pour les phrases 1 (/valmõ vulɛ vvar lə vɔl/) et 3 (/mamã e mami mani ma mẽ/) dans chaque position des essais expérimentaux dans l'expérience I .....	93
Figure 21 Moyenne et écart-type de la PLV pour chaque type de phrase dans chaque position des essais expérimentaux pour les locutrices dans la « condition uniquement DAF ») dans l'expérience II.....	95
Figure 22 Moyenne et écart-type de la PLV dans chaque SA pour les locutrices dans la « condition uniquement DAF » dans l'expérience II, respectivement pour chaque catégorie d'analyse .....	97
Figure 23 PLV moyenne de chaque type de phrase des répétitions en position 1 (axe X) et en position 3 (axe Y) pour les locutrices dans la « condition uniquement DAF » dans l'expérience II. ....	99
Figure 24 PLV moyenne de chaque type de phrase produite en position 1 pour les locutrices dans la « condition uniquement DAF » dans essais sans DAF (axe X) et dans les essais avec DAF à 120 ms (axe Y) dans l'expérience II.....	100
Figure 25 Distribution du nombre de répétitions de phrases contenant différents nombres de pics syllabiques (sur les colonnes) et de pics supra-syllabiques (sur l'axe des X) identifiés grâce à notre méthode d'analyse des modulations d'amplitude (cf. chapitre III 3.2), respectivement pour chaque condition du DAF (en colonnes), concernant les locutrices dans la « condition uniquement DAF » de l'expérience II.....	102
Figure 26 PLV de phrase à chaque niveau du DAF (axe X) dans les répétitions de phrases où différents nombres de pics syllabiques (lignes) et supra-syllabiques (colonnes) ont été trouvés par notre méthode d'analyse des modulations d'amplitude (cf. chapitre III 3.2), pour les locutrices dans la « condition uniquement DAF » de l'expérience II .....	103
Figure 27 Modulations d'amplitude pour la phrase « Jacqueline gère le jour » (/zaklin ʒɛʁ lə ʒuʁ/) .....	104
Figure 28 F0 médiane des voyelles (accentuables ou non accentuables) produites à différents niveaux du décalF0, respectivement pour chaque locutrice, dans les essais sans DAF dans l'expérience II (condition perturbations multiples).....	118
Figure 29 Moyenne et écart-type de la durée des voyelles dans le <b>SA final</b> pour chaque niveau du DAF et chaque niveau du <b>décalF0 positif</b> dans l'expérience II. ....	121
Figure 30 Moyenne et écart-type de la durée des voyelles dans chaque SA pour chaque niveau du DAF et chaque groupe de locutrices, dans l'expérience II.....	123
Figure 31 PLV du <b>SA initial</b> et du <b>SA final</b> dans les répétitions de chaque type de phrase en position 1 dans un essai, pour chaque niveau du DAF (axe X) et chaque niveau du décalF0 (couleur), pour les locutrices ayant suivi le décalF0 <b>positif</b> , dans l'expérience II .....	126
Figure 32 PLV du <b>SA initial</b> et du <b>SA final</b> dans les répétitions des phrases en position 1 dans un essai, pour chaque niveau du DAF (axe X) et chaque niveau du décalF0 (couleur), pour les locutrices ayant suivi le décalF0 <b>néгатif</b> , dans l'expérience II.....	127
Figure 33 PLV du <b>SA initial</b> et du <b>SA final</b> dans les répétitions des phrases en position 1 dans un essai, pour chaque niveau du DAF (axe X) et chaque groupe de locutrices (couleur), dans l'expérience II.....	130
Figure 34 Moyenne et écart-type de la f0 médiane des voyelles dans le <b>SA final</b> dans chaque type de phrase (lignes) pour chaque niveau du DAF (axe X) et chaque groupe de locutrices (colonne), dans l'expérience II.....	132

---

Figure 35 Moyenne et écart-type de l'intensité médiane des voyelles dans le **SA initial** de chaque type de phrase (ligne) pour chaque niveau du DAF (axe X) et chaque groupe de locutrices (colonne), dans l'expérience II.....133

Figure 36 Moyenne et écart-type de l'intensité médiane des voyelles dans le **SA final** de chaque type de phrase (ligne) pour chaque niveau du DAF (axe X) et chaque groupe de locutrices (colonne), dans l'expérience II.....134

---

## TABLE DES TABLEAUX

Tableau 1 Résumé des résultats du modèle linéaire mixte portant sur l'effet du DAF sur la durée des voyelles dans différentes positions de la <b>phrase 1</b> dans l'expérience I (cf. chapitre IV 1.1)..	165
Tableau 2 Résumé des résultats du modèle linéaire mixte portant sur l'effet du DAF sur la durée des voyelles dans différentes positions de la <b>phrase 3</b> dans l'expérience I (cf. chapitre IV 1.1) .	166
Tableau 3 Résumé des résultats du modèle linéaire mixte portant sur l'effet du DAF sur la durée des voyelles dans différentes positions pour le groupe des locutrices dans la « condition uniquement DAF) dans l'expérience II (cf. chapitre IV 1.2).....	167
Tableau 4 Résumé des résultats du modèle linéaire mixte portant sur l'effet du DAF sur le ratio entre la durée totale des voyelles accentuables et non accentuables (ratioDur) dans l'expérience I (cf. chapitre IV 2.1).....	168
Tableau 5 Résumé des résultats des deux modèles linéaires mixtes portant sur l'effet du DAF sur le ratio entre la durée de la voyelle accentuée et celle de la voyelle non-accentuée dans chaque SA (V2/V1 et V5/V4) pour les locutrices dans la « condition uniquement DAF » dans l'expérience II, respectivement pour chaque catégorie d'analyse (cf. chapitre IV 2.2).....	168
Tableau 6 Résumé des résultats du modèle linéaire mixte portant sur l'effet du DAF sur la PLV dans l'expérience I (cf. chapitre V section 1).....	169
Tableau 7 Résumé des résultats du modèle linéaire mixte portant sur l'effet du DAF sur la PLV d'une phrase entière pour les locutrices dans la « condition uniquement DAF ») dans l'expérience II (cf. chapitre V 2.1).....	169
Tableau 8 Résumé des résultats des deux modèles linéaires mixtes portant sur l'effet du DAF sur la PLV de chaque SA pour les locutrices dans la « condition uniquement DAF » dans l'expérience II, respectivement pour chaque catégorie d'analyse (cf. chapitre V 2.2).....	170
Tableau 9 Récapitulatif des 20 modèles linéaires portant sur l'effet décalF0, conduits individuellement pour chaque locutrice (ex. Loc01 = locutrice 01) ayant subi le décalF0 dans l'expérience II (cf. chapitre VI section 1).....	170
Tableau 10 Résumé des résultats des 20 modèles linéaires portant sur l'effet du décalF0, conduits individuellement pour chaque locutrice ayant subi le décalF0 dans l'expérience II (cf. chapitre VI section 1) .....	171
Tableau 11 Résumé des résultats des 12 modèles linéaires mixtes comparant l'effet du DAF sur la durée des voyelles dans chaque SA de chaque type de phrase entre les différentes conditions du décalF0 pour les locutrices ayant suivi le décalage dans l'expérience II (cf. chapitre VI 2.1.1)....	172
Tableau 12 Résumé des résultats des six modèles linéaires mixtes comparant l'effet du DAF sur la durée des voyelles dans chaque SA de chaque type de phrase entre les locutrices dans la « condition uniquement DAF » et les locutrices dans la « condition perturbations multiples » dans l'expérience II (cf. chapitre VI 2.1.2).....	175
Tableau 13 Résumé des résultats des six modèles linéaires mixtes comparant l'effet du DAF sur la PLV (d'une phrase entière ou de chaque SA) entre les différentes conditions du décalF0 dans l'expérience II (cf. chapitre VI 2.2.1) .....	177
Tableau 14 Résumé des résultats des trois modèles linéaires mixtes comparant l'effet du DAF sur la PLV (d'une phrase entière ou de chaque SA) entre les locutrices dans la « condition uniquement	

---

DAF » et les locutrices dans la « condition perturbations multiples » dans l'expérience II (cf. chapitre VI 2.2.2) .....	177
Tableau 15 Résumé des résultats des six modèles linéaires mixtes comparant l'effet du DAF sur la f0 médiane des voyelles dans chaque SA (initial ou final) entre le groupe des locutrices dans la « condition uniquement DAF » et les locutrices dans la « condition perturbations multiples » dans l'expérience II (cf. chapitre VI 2.3.1).....	178
Tableau 16 Résumé des résultats des six modèles linéaires mixtes comparant l'effet du DAF sur l'intensité médiane des voyelles dans chaque SA (initial ou final) entre les locutrices dans la « condition uniquement DAF » et les locutrices dans la « condition perturbations multiples » dans l'expérience II (cf. chapitre VI 2.3.2).....	179

---

# GLOSSAIRE

Dans ce glossaire, nous présentons les définitions de certains termes utilisés dans le cadre de cette thèse.

## 1. Syntagme accentuel

Dans le contexte de cette thèse, nous considérons le syntagme accentuel comme l'unité prosodique où s'applique l'accent primaire, en suivant la définition de Di Cristo et Hirst ou Jun et Fougeron (2000). Un syntagme accentuel est constitué d'un mot lexical portant l'accent primaire et d'une ou plusieurs syllabes qui le précèdent et en dépendent (par exemple, « une jolie maison »).

## 2. Groupement rythmique

Le groupement rythmique se réfère à l'organisation des éléments linguistiques en groupes, et cette organisation est influencée par la structure prosodique. Dans le contexte de cette thèse, notre intérêt se porte particulièrement sur le groupement syllabique, c'est-à-dire, l'organisation des syllabes entre elles.

## 3. Modulation d'amplitude

La modulation d'amplitude est une méthode permettant d'extraire les enveloppes d'amplitude, qui représentent les variations d'amplitude à une certaine fréquence dans le signal acoustique. Les signaux extraits prennent la forme de sinusoides.

## 4. Composante syllabique

Une composante syllabique est le signal oscillatoire extrait à l'aide de la méthode de la modulation d'amplitude, correspondant aux variations dues à la production des syllabes dans le signal acoustique.

## 5. Composante supra-syllabique

Une composante supra-syllabique est le signal oscillatoire extrait à l'aide de la méthode de la modulation d'amplitude, correspondant aux variations plus lentes résultant de la production des syllabes plus proéminentes dans le signal acoustique. Les syllabes proéminentes font référence aux syllabes présentant une saillance d'énergie.

## 6. Coordination entre les composantes rythmiques

Une forte coordination entre les composantes rythmiques correspond à une stabilité élevée de la relation temporelle entre la composante syllabique et la composante supra-syllabique. Le degré

---

de cette coordination est mesuré à l'aide de la méthode de *Phase Locking Value* (PLV), qui quantifie la variabilité du décalage entre les deux signaux par rapport à la durée de leurs cycles dans chaque fenêtre temporelle considérée.

## **7. Réorganisation rythmique**

Dans le contexte de cette thèse, une réorganisation rythmique fait référence à une variation du degré de coordination entre les composantes rythmiques.

## **8. Intégration sensorimotrice**

Le système de production de la parole est en constante évaluation de l'état du système en comparant les informations sensorielles résultant de nos propres articulations (les retours sensoriels) aux conséquences sensorielles souhaitées. Cette comparaison permet d'ajuster en temps réel les activités motrices pour s'adapter aux perturbations imprévues ou à différentes situations de la production de la parole.

## **9. Perturbation du retour auditif**

La perturbation du retour auditif est un paradigme expérimental dans lequel la voix des locuteurs est artificiellement modifiée avant d'être renvoyée aux locuteurs à travers un casque. Cette modification artificielle peut consister en un retard du moment où les locuteurs entendent leur voix ou en un décalage (positif ou négatif) de la fréquence fondamentale ( $f_0$ ) de leur voix.

## **10. Sens d'agentivité**

Le sens d'agentivité est la perception du contrôle sur une action générée par soi-même. Il est lié au sentiment d'être l'auteur qui a provoqué une action et au sentiment d'être responsable des conséquences de cette action.

# INTRODUCTION GENERALE

## 1. Contexte

Parler est la capacité humaine de construire des séquences sonores pour exprimer ses intentions. Dans la production de ces séquences sonores, il est nécessaire de contrôler des dizaines de muscles afin d'orchestrer les mouvements de plusieurs articulateurs. Cependant, l'exécution de cette tâche se trouve confrontée à au moins deux facteurs qui la complexifient. D'un côté, la production de la parole est soumise à des contraintes matérielles. Plus précisément, il est indispensable de réguler les interactions entre divers paramètres neurophysiologiques, qui ne sont que partiellement indépendants les uns des autres. Par exemple, les phénomènes de coarticulation, où les sons d'une séquence sont influencés par les sons qui les entourent, sont fréquemment observés. De l'autre côté, la production de la parole est également contrainte par des facteurs linguistiques, lesquels peuvent varier d'une langue à l'autre. Par exemple, l'ajustement des interactions mentionnées précédemment doit prendre en compte toutes sortes de contraintes définies à différents niveaux de la hiérarchie linguistique. Ainsi, la coarticulation dépend, par exemple, de la structure prosodique spécifique à chaque langue.

Malgré la complexité de production de la parole, le contrôle moteur impliqué doit être suffisamment flexible pour faire face à divers types de « perturbations » dont l'occurrence n'est pas prédictible. Ces perturbations peuvent découler de plusieurs sources : 1) la complexité inhérente au système moteur est à l'origine du bruit interne au système et des perturbations internes générées ; 2) les changements environnementaux tels que l'apparition soudaine d'un bruit de fond ou d'événements imprévus qui affectent nos intentions communicatives peuvent nous amener à restructurer nos plans moteurs impliqués dans la production de la parole ; 3) les facteurs liés à la nature interactive de la communication, tels que la présence d'un interlocuteur, peuvent également nous amener à replanifier nos commandes motrices.

D'un côté, pour réduire la complexité de la production de la parole, les mouvements des articulateurs sont coordonnés pour atteindre des objectifs linguistiques communs. Cette

coordination engendre un rythme coordinatif (Poeppel et Assaneo, 2020), qui, bien que contraint par les facteurs neurophysiologiques et linguistiques mentionnés précédemment, contribue à simplifier la dimension du contrôle moteur et, en retour, à stabiliser la coordination des mouvements des articulateurs. De l'autre côté, pour s'adapter de manière flexible aux perturbations imprévues ou aux changements dans l'environnement, le système de contrôle moteur des locuteurs suit les retours sensoriels résultant des commandes motrices générées (c'est-à-dire les conséquences sensorielles des articulations) lors de la production de la parole et les compare aux attentes. Lorsque des observations inattendues surviennent, le système moteur émet immédiatement des commandes de compensation afin de réduire l'écart entre les attentes et les observations. Un exemple concret de cette adaptation est observé chez les fumeurs, qui ajustent la forme de leur bouche et la position de leur langue pour produire la parole attendue lorsqu'ils parlent avec une cigarette à la bouche.

La prosodie joue un rôle essentiel dans la modulation du rythme de la parole en attribuant de la saillance à certaines syllabes dans les énoncés par le biais de l'accentuation, ce qui entraîne le groupement des syllabes. Plusieurs études ont tenté d'explorer le rôle de l'accentuation dans les processus de suivi des retours sensoriels par le contrôle moteur (Kalveram et Jancke, 1989 ; Oschkinat et Hoole, 2022). Cependant, très peu d'études se sont intéressées à l'adaptation du groupement rythmique (ex. le groupement des syllabes) au sein de ces processus.

De plus, lorsque les locuteurs sont impliqués dans une interaction verbale, ils s'adaptent rapidement aux caractéristiques de la parole de leurs interlocuteurs et ajustent leur propre discours en conséquence. Dans ce contexte, les différences entre les caractéristiques de la parole entendue et celle prédite en fonction des habitudes articulatoires des locuteurs ne déclenchent pas de réponses compensatrices du système moteur. En effet, de telles corrections accentueraient les divergences entre les caractéristiques de la parole des interlocuteurs, entravant ainsi la fluidité de la communication entre eux. Au contraire, les locuteurs mettent à jour leurs attentes en fonction des caractéristiques de la parole de leurs interlocuteurs, ce qui les conduit à imiter les caractéristiques de la parole de leur interlocuteur (Pickering et Garrod, 2004). Par exemple, il est courant que les locuteurs suivent l'accent de leur interlocuteur ou adoptent sa façon de parler (ex. Pardo, 2006).

Par conséquent, il est essentiel pour les locuteurs de pouvoir différencier leur propre voix de celle d'une autre personne afin de déterminer comment ajuster leurs commandes motrices. Les facteurs d'agentivité et de sens d'agentivité jouent donc un rôle crucial dans la production de la parole. L'agentivité fait référence à la capacité d'un individu à contrôler les actions résultant de ses

propres intentions, tandis que le sens d'agentivité représente la perception subjective d'un individu de sa propre agentivité, incluant la perception qu'un événement est le résultat de ses propres intentions. Cependant, il a été démontré que la relation entre l'agentivité et le sens d'agentivité peut varier d'un locuteur à l'autre pendant la production de la parole. Par exemple, dans les études où les locuteurs entendaient la fréquence fondamentale ( $f_0$ ) de leur propre voix manipulée (ex. en augmentant la  $f_0$  de leur voix) et retournée à travers un casque, la plupart des locuteurs compensaient la manipulation de la  $f_0$  en modifiant leur propre  $f_0$  dans la direction opposée à la manipulation (Burnett et al., 1998 ; Hain et al., 2000 ; Larson et al., 2007 ; Behroozmand et al., 2012 ; Franken et al., 2018). Cependant, certains locuteurs ont parfois suivi la manipulation au lieu de la compenser. Étant donné que le degré d'agentivité était manipulé (réduit) de la même manière pour les locuteurs ayant présenté un comportement imitatif et pour ceux ayant présenté un comportement compensateur, il est raisonnable de supposer que c'est le sens d'agentivité qui est affecté différemment chez les deux groupes de locuteurs. Ainsi, le sens d'agentivité des locuteurs est variable et flexible. Il est donc intéressant d'explorer comment le sens d'agentivité peut être manipulé et comment l'altération du sens d'agentivité peut influencer la production de la parole.

L'objectif principal de cette thèse est d'étudier la production de la parole dans une perspective intégrative, en prenant en compte à la fois la complexité et la flexibilité de la production de la parole. Plus précisément, nous avons examiné l'interaction entre la prosodie et les processus du contrôle sensorimoteur, ainsi que l'influence de cette interaction sur l'organisation rythmique (ex. groupement des syllabes). Nous avons également examiné comment le sens d'agentivité module la production de la parole et son interaction avec la prosodie en production de la parole.

## 2. Structure de la thèse

Ce manuscrit est composé de sept chapitres. Il débutera par deux chapitres visant à établir le contexte de recherche. Le **chapitre I** révisera la complexité inhérente à la production de la parole et examinera comment le rythme est généré et organisé pour réduire cette complexité. Nous discuterons également de la modulation du rythme par la prosodie et proposerons une nouvelle perspective pour l'étude de l'organisation rythmique, en considérant le rythme comme un système

coordonatif. Le **chapitre II** présentera le mécanisme du contrôle sensorimoteur qui permet au système moteur de s'adapter de manière flexible à différentes situations de la production de la parole. Dans ce même chapitre, nous discuterons également du rôle de deux facteurs, la prosodie et le sens d'agentivité, dans les processus du contrôle sensorimoteur, en nous appuyant sur les études existantes. Nous ouvrirons la perspective de recherche sur l'adaptation de l'organisation rythmique aux conditions du contrôle sensorimoteur.

Une fois le contexte établi, le **chapitre III** présentera nos deux questions de recherche et les hypothèses correspondantes. Nous y détaillerons également les deux expériences que nous avons menées en utilisant le paradigme des perturbations du retour auditif, à savoir le retour auditif retardé et le décalage continu de la  $f_0$  du retour auditif, pour explorer nos hypothèses. Les méthodes d'analyse seront également explicitées dans le cadre de ce chapitre.

Les trois chapitres suivants exposeront les analyses et leurs résultats, chacun de ces chapitres examinant une ou deux hypothèses spécifiques que nous avons formulées. Dans le **chapitre IV**, nous mettrons en avant les analyses axées principalement sur l'étude du rôle de l'accentuation dans les processus du contrôle sensorimoteur. Le **chapitre V** sera consacré aux analyses portant sur l'adaptation de l'organisation rythmique dans le contexte de la perturbation temporelle du retour auditif (le retour auditif retardé). Enfin, le **chapitre VI** se concentrera sur les analyses examinant le rôle du sens d'agentivité et son interaction avec la prosodie dans le contrôle sensorimoteur. À la fin de chaque chapitre d'analyses, nous fournirons une discussion des résultats.

Nous concluons ce manuscrit par une synthèse générale de cette thèse et une présentation des perspectives de recherche ouvertes par cette étude dans le **chapitre VII**.

# CHAPITRE I : COMPLEXITE DE LA PRODUCTION DE LA PAROLE

## Sommaire du chapitre

<b>1. Groupement rythmique réduisant la complexité du contrôle moteur</b> .....	<b>6</b>
1.1 Articulateurs travaillant en synergie .....	6
1.2 Émergence du rythme .....	7
1.3 Rythme syllabique .....	9
1.4 Hiérarchisation du groupement rythmique .....	10
<b>2. Lien et différences entre la prosodie et le rythme</b> .....	<b>12</b>
2.1 Accentuation, métrique et unités prosodiques .....	13
2.2 Unités prosodiques du français .....	14
2.2.1 Application de l'accentuation en français .....	14
2.2.2 Système métrique du français .....	15
2.2.3 Caractéristiques de l'accentuation du français .....	16
2.3 Formalisation des phénomènes métriques .....	17
2.4 Rythme dérivant de la prosodie tout en s'en distinguant .....	18
2.5 Mesurer des régularités rythmiques .....	19
<b>3. Prosodie en tant que mécanisme stabilisant du contrôle moteur</b> .....	<b>21</b>
3.1 Rythme en tant que système coordinatif .....	21
3.2 Coordination entre les composantes rythmiques à différentes échelles temporelles .....	22
3.2.1 Composante syllabique et composante supra-syllabique .....	22
3.2.2 Coordination entre la composante syllabique et la composante supra-syllabique .....	23
<b>4. Résumé du chapitre</b> .....	<b>25</b>

La production de la parole est une tâche complexe qui englobe divers processus visant à atteindre nos objectifs linguistiques en mobilisant les mouvements des articulateurs pour générer des signaux acoustiques. Ces processus intègrent des éléments à la fois abstraits (ex. la planification des commandes motrices en fonction des objectifs linguistiques), et matériels (ex. la coordination des mouvements des articulateurs pour exécuter les commandes motrices). Les processus inhérents à la production de la parole englobent également des mécanismes de suivi et d'adaptation, qui garantissent la stabilité et la flexibilité du processus de production de la parole. Ces mécanismes seront détaillés dans le chapitre [II](#).

Ce premier chapitre sera d'abord consacré à l'exploration de l'émergence du groupement rythmique, un processus qui contribue à atténuer les contraintes matérielles liées à la complexité

inhérente du contrôle moteur au cours de la production de la parole (section 1). Ensuite, nous analyserons la manière dont la structure prosodique module différents niveaux du groupement rythmique (section 2). Enfin, en nous appuyant sur le concept de groupement et la hiérarchisation du rythme, nous développerons la perspective selon laquelle le rythme peut être envisagé comme un système coordinatif, dont les interactions entre ses composantes varient en fonction des langues et des contextes de production de la parole (section 3).

## **1. Groupement rythmique réduisant la complexité du contrôle moteur**

### **1.1 Articulateurs travaillant en synergie**

Plusieurs articulateurs, tels que la langue, les lèvres et la mâchoire, interviennent dans le processus de production des sons de la parole. Cependant, l'orchestration de ces articulateurs représente une tâche d'une grande complexité. Par exemple, nous observons fréquemment des phénomènes tels que la coarticulation, où la réalisation physique d'un son est influencée par les sons environnants dans une séquence sonore (cf. par exemple De Jong et al., 1993). En d'autres termes, les sons ne sont pas traités comme des unités linéaires dans le temps. Par conséquent, les mouvements des articulateurs ne peuvent non plus être organisés de manière linéaire dans le temps. En réalité, pour atteindre un objectif acoustique donné, la cinématique (ex. la position, la vitesse et l'accélération), n'est pas définie de manière indépendante pour chaque articulateur. Les mouvements des articulateurs sont coordonnés de manière synergique, formant ainsi une action coordonnée au sein du conduit vocal (Abbs et Gracco, 1984 ; Browman et Goldstein, 1992). C'est donc la relation dynamique entre les mouvements des articulateurs qui joue un rôle crucial dans la production de la parole.

Ce phénomène de synergie est décrit dans le domaine de la phonologie articuloire (Browman et Goldstein, 1989 ; Browman et Goldstein, 1992 ; Goldstein et Fowler, 2003 ; Fougeron, 2005) comme le groupement et la coordination de plusieurs articulateurs en gestes articuloires, qui se manifestent par des constriction à des endroits spécifiques dans le conduit vocal. Chaque constriction contribue à ajuster l'ouverture du conduit vocal. Prenons l'exemple d'un geste « labial

» (le geste nécessaire pour produire, par exemple, le son bilabial [p]). Dans ce cas, les trois articulateurs impliqués (la lèvre inférieure, la lèvre supérieure et la mâchoire) se coordonnent pour exécuter une action dans le conduit vocal.

## 1.2 Émergence du rythme

La coordination des articulateurs induit une corrélation significative dans leurs mouvements. Pour illustrer ce phénomène, la Figure 1, adaptée de l'étude menée par Poeppel et Assaneo (2020), offre un aperçu des mouvements verticaux de trois articulateurs (la lèvre inférieure, la pointe de la langue et l'incisive inférieure), pendant la production de la phrase en anglais : « *Before Thursday's exam, review every formula* ». Cette illustration inclut également l'oscillation résultant de la multiplication de la racine carrée des mesures de ces mouvements articulatoires (visible sous la forme d'une oscillation en bleu clair), ainsi que le signal acoustique associé à la phrase et son enveloppe d'amplitude (présentés sous la forme d'une oscillation en rouge). Nous observons que les mouvements des trois articulateurs sont généralement corrélés, ce qui donne naissance à un rythme de coordination (cf. l'oscillation en bleu clair). Ce rythme de coordination reflète la dynamique coordonnée des trois articulateurs. Nous pouvons apprécier la relation entre le rythme collectif et le comportement individuel de chaque articulateur dans la Figure 1 en observant que chaque transition dans le rythme collectif (indiquée par une bulle rouge) est synchronisée avec au moins une transition dans le mouvement d'un des articulateurs (indiquée par une bulle bleue).

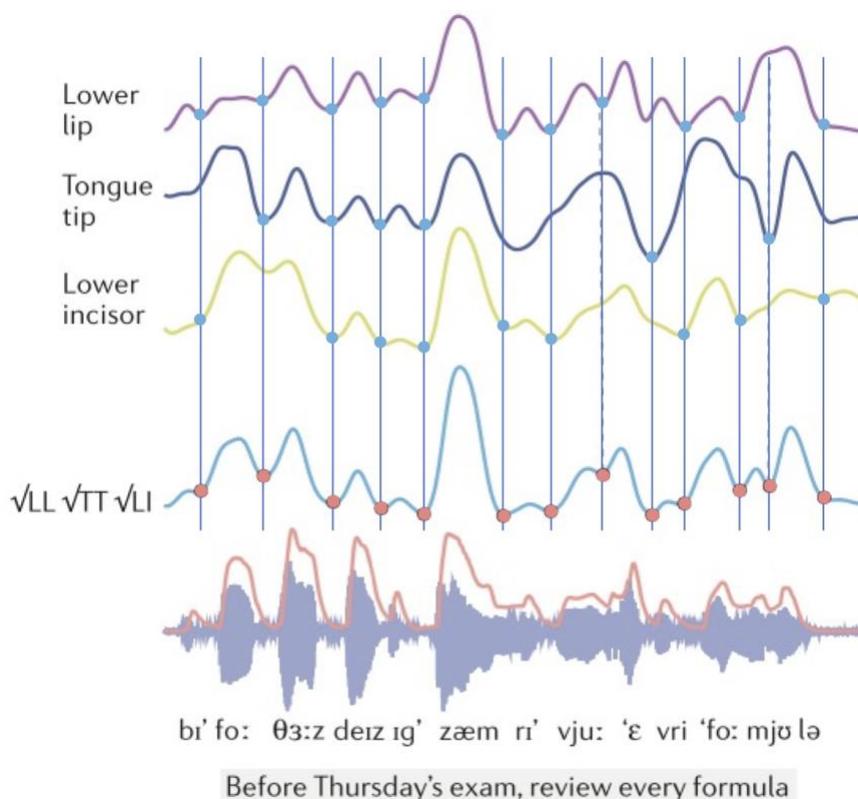
Ainsi, l'interruption du rythme d'un des articulateurs n'interrompt pas le rythme coordonné (Poeppel et Assaneo, 2000). Cette caractéristique permet à chaque articulateur de compenser les perturbations survenues chez les autres articulateurs, sans nécessiter une replanification des commandes motrices pour atteindre l'objectif. Par exemple, dans l'étude de Kelso et ses collègues (1984), l'application soudaine d'une charge constante pendant le mouvement ascendant de la mâchoire lors de la production du /b/ final dans la syllabe /bæb/ a conduit les participants à compenser cette perturbation quasiment instantanément au niveau des lèvres supérieure et inférieure. Cette compensation rapide démontre qu'il est quasiment impossible de replanifier les commandes motrices en réponse à une perturbation. En d'autres termes, le système moteur s'autoorganise pour faire face aux perturbations au niveau périphérique, sans nécessiter de calculs

complexes par le système de contrôle central. Ainsi, la coordination des articulateurs favorise les réponses aux perturbations inattendues lors de la production de la parole.

Ainsi, en groupant les mouvements des articulateurs en gestes articulatoires, le système de contrôle moteur parvient à réduire ses dimensions de contrôle (Poeppel et Assaneo, 2020). En d'autres termes, au lieu de contrôler individuellement chaque articulateur, le système de contrôle moteur orchestre la coordination entre les mouvements de ces articulateurs (Saltzman et Kelso, 1987 ; Browman et Goldstein, 1992). Cette constatation a également été corroborée par des études sur les activités neurales, telles que celle menée par Chartier et al (2018), qui ont démontré que le cortex moteur encode les mouvements coordonnés des articulateurs nécessaires pour atteindre une constriction du conduit vocal, plutôt que d'encoder de manière indépendante des mouvements de chaque articulateur.

Figure 1 Illustration de l'alignement des régularités rythmiques entre les articulateurs, ainsi qu'entre le domaine articulatoire et le domaine acoustique, adaptée de Poeppel et Assaneo (2020)

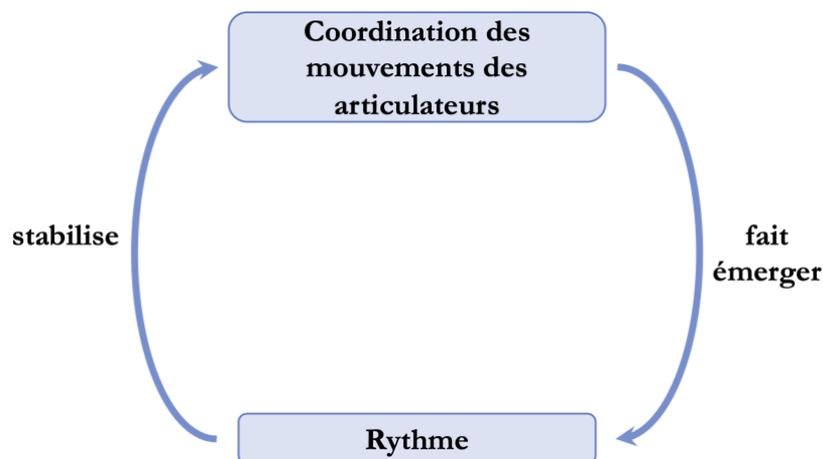
Les trois oscillations supérieures représentent le déplacement vertical de la lèvre inférieure (LL), de la pointe de la langue (TT) et de l'incisive inférieure (LI) pour un sujet pendant la production de la phrase en anglais indiquée en bas de la figure. L'oscillation en bleu clair illustre la multiplication de la racine carrée des mesures articulatoires. Le panneau le plus bas affiche le signal acoustique de la phrase avec son enveloppe d'amplitude en oscillation rouge.



De ce fait, le groupement et la coordination des mouvements des articulateurs contribuent à simplifier la complexité de la production de la parole en générant un rythme de plus haut niveau

qui dirige le système de contrôle moteur. Ce rythme supérieur réduit le degré de liberté de chaque sous-système (ex. les mouvements de chaque articulateur). En d'autres termes, la coordination entre les mouvements des articulateurs constitue la fondation du rythme de la parole (selon Poeppel et Assaneo, 2020). Une fois que le rythme de la parole émane des mouvements des articulateurs, il contribue à la régulation de ces mouvements (cf. Figure 2).

Figure 2 Illustration de la relation entre la production de la parole et le rythme de la parole



### 1.3 Rythme syllabique

Dans la Figure 1, nous pouvons non seulement observer un quasi-alignement des mouvements des articulateurs, mais également une cohérence entre le rythme coordinatif (l'oscillation en bleu claire) et l'oscillation de l'enveloppe d'amplitude du signal acoustique de la phrase produite. L'oscillation de l'enveloppe d'amplitude représente les variations d'amplitude du signal acoustique. Étant donné que les minima locaux dans l'oscillation de l'enveloppe d'amplitude définissent pratiquement les frontières syllabiques (Tilsen et Arvaniti, 2013 ; Mermelstein, 1975), nous pouvons en déduire que la coordination des mouvements des articulateurs donne naissance à un rythme syllabique. En effet, l'étude de Goldstein (2019) suggère que les mouvements articulatoires exhibent un rythme inhérent. En d'autres termes, ces mouvements articulatoires se disposent en alternance entre des périodes de changement rapide et des périodes de changement modéré. De surcroît, cette alternance de périodes est corrélée au rythme syllabique, comportant une ou deux alternances au cours de la production de chaque syllabe.

Le rythme syllabique joue un rôle significatif dans l'intelligibilité de la parole (Ghitza et Greenberg, 2009). Il est remarquable que ce rythme demeure dans une plage de fréquences relativement restreinte, soit de 2,5 à 8 syllabes par seconde (Poeppel et Assaneo, 2020). Cette constance est observée malgré les variations linguistiques, les différences entre locuteurs, les variations de vitesse d'élocution et les conditions de production de la parole. Des études (Doelling et al., 2014 ; Assaneo et Poeppel, 2018 ; Ghitza et Greenberg, 2009) ont mis en évidence que le cortex auditif suit les enveloppes de la parole, et cette tendance à suivre est particulièrement renforcée pour la plage de fréquences correspondant au rythme syllabique dans les enveloppes des signaux acoustiques. Les régularités rythmiques des enveloppes des signaux acoustiques permettent au cortex auditif de segmenter le signal continu de la parole en unités appropriées pour faciliter l'encodage (Giraud et Poeppel, 2012).

Le rythme syllabique est donc un phénomène complexe. D'un côté, comme décrit précédemment, il émerge de la coordination de multiples articulateurs. D'un autre côté, de nombreuses contraintes, qu'elles soient d'ordre biomécanique, neurophysiologique ou cognitif (MacNeilage, 1998 ; Davis et MacNeilage, 2004 ; Di Cristo, 2003 ; Assaneo et Poeppel, 2018), influencent le contrôle de ces articulateurs. Par ailleurs, il convient de noter que ce rythme est également influencé par des facteurs linguistiques, que nous explorerons plus en détail dans la suite de ce chapitre.

La syllabe ne se limite donc pas à une simple unité phonologique, elle joue également un rôle essentiel dans la planification motrice et la régulation de la production de la parole. En phonologie articulatoire, les gestes articulatoires, qui sont les unités phonologiques minimales, sont organisés en syllabes plutôt qu'en segments (Browman et Goldstein, 1990). De plus, dans le cadre du modèle DIVA (*Directions into Velocities of Articulators*), un réseau neuronal élaboré par Guenther et ses collaborateurs (2006), qui permet d'expliquer de nombreux phénomènes liés à la production et à l'acquisition de la parole, la syllabe est fréquemment utilisée en tant qu'unité de base pour activer les commandes motrices.

## 1.4 Hiérarchisation du groupement rythmique

Dans la section [1.3](#), nous avons observé que la coordination (le groupement) des gestes articulatoires favorise l'apparition du rythme syllabique. Revenons brièvement à l'exemple présenté

dans la Figure 1. Le signal acoustique et de son enveloppe d'amplitude en bas de la Figure 1 révèlent que les syllabes elles-mêmes sont également groupées en unités plus grandes. Par conséquent, il est raisonnable de déduire que le groupement des éléments dans la production de la parole est organisé de manière hiérarchique. En d'autres termes, le rythme de la parole suit une organisation hiérarchique.

La formation d'un groupe nécessite inévitablement une tête d'unité (Delais-Roussarie, 2022). La tête d'une unité est généralement perçue comme étant plus saillante que les autres éléments au sein de cette unité. Conformément aux principes de la psychologie gestaltiste (Köhler, 1967 ; Di Cristo, 2003), en ce qui concerne la perception de la parole, la présence de saillance au sein d'un flux de parole induit une rupture dans la perception, engendrant ainsi des groupements perceptifs.

La tête d'une unité syllabique, formée par le groupement des gestes articulatoires, est le noyau syllabique. Ce noyau syllabique correspond généralement à la voyelle dans une syllabe, dans les langues qui requièrent la présence d'une voyelle dans chaque syllabe, comme le français. Le noyau syllabique est généralement produit avec une énergie accrue au sein de la syllabe. Ainsi, dans la Figure 1, les maxima de l'enveloppe d'amplitude doivent correspondre aux noyaux syllabiques.

D'autre part, la tête d'une unité formée par le groupement des syllabes doit être la syllabe la plus saillante au sein de cette unité. Cette syllabe la plus saillante est souvent une syllabe accentuée, ayant un rôle phonologique important (cf. chapitre I 2.1). Ainsi, le groupement rythmique (c'est-à-dire le groupement des éléments de la production de la parole en unités plus grandes) est étroitement lié à l'organisation de l'accentuation. L'accentuation détermine une syllabe comme plus proéminente que celles qui l'entourent et constitue un paramètre essentiel de la prosodie de la parole. L'accentuation est organisée de manière hiérarchique. De plus, chaque langue possède ses propres caractéristiques prosodiques. Ainsi, la prosodie demeure une source cruciale de contraintes linguistiques dans l'organisation du rythme. Nous examinerons plus en détail le lien entre la prosodie et le rythme de la parole dans la section 2 de ce chapitre.

Cette définition du groupement rythmique comporte trois implications :

1) Le rythme ne se réduit pas à un simple phénomène de régularité temporelle. Il englobe la formation de groupements entre les éléments pour constituer des unités plus grandes et la relation entre ces éléments au sein de ces unités (Fraisse, 1956 ; Fónagy et Fónagy, 1983 ; Wioland, 1984 ; cité par Di Cristo et Hirst, 1993). Cette relation est définie par le degré de saillance de chaque élément.

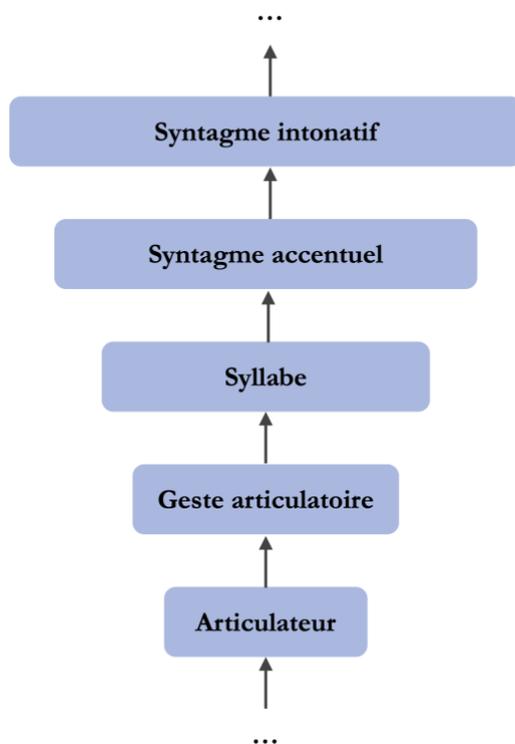
2) En principe, chaque niveau prosodique pourrait constituer un degré de groupement rythmique (cf. Figure 3 et chapitre I section 2).

3) Plusieurs paramètres acoustiques liés à la prosodie, tels que la durée, l'intensité et la  $f_0$ , peuvent influencer le groupement rythmique (Arvaniti, 2009).

Au sein de cette thèse, notre focalisation prédominante se concentre sur l'analyse du groupement syllabique, c'est-à-dire la manière dont plusieurs syllabes se combinent les unes avec les autres.

Figure 3 Illustration des niveaux du groupement rythmique

*L'élément contenu dans chaque boîte bleue correspond au groupement des éléments contenus dans la boîte immédiatement située en dessous. Pour obtenir la définition de l'élément « geste articulatoire », voir la section 1.1 de ce chapitre. Pour accéder à la définition des éléments situés au-dessus de l'élément « syllabe », voir la section 2.2.1.*



## 2. Lien et différences entre la prosodie et le rythme

La prosodie découpe le flux de parole en éléments hiérarchiquement organisés, facilitant ainsi la perception et l'encodage de l'information transmise par les locuteurs aux auditeurs. La prosodie se présente comme un candidat idéal pour moduler le rythme de la parole. Plus précisément, la

prosodie est un paramètre abstrait qui joue un rôle de structuration dans la production de la parole, en appliquant des règles préétablies à l'organisation des éléments de la production de la parole. Le rythme, par conséquent, peut être considéré comme la matérialisation de la prosodie. Toutefois, nous ne pouvons pas nous attendre à ce que les patrons rythmiques de la parole soient entièrement prédictibles sur la base des règles qui décrivent le système prosodique, car le rythme de surface résulte de l'interaction de diverses contraintes, notamment d'origine biomécanique, cognitive (comme brièvement mentionné dans la section [1](#) de ce chapitre), ou de divers facteurs non linguistiques (ex. la situation de la production de la parole).

Nous entamons cette section par une description de la structure prosodique, qui englobe l'accentuation en tant que paramètre crucial dans la formation de cette structure, la métrique qui organise l'accentuation, ainsi que les unités prosodiques à différents niveaux de la structure prosodique. Nous portons une attention particulière aux unités prosodiques du français, qui est la langue d'étude de cette thèse. Nous aborderons également la distinction entre la prosodie et le rythme.

## **2.1 Accentuation, métrique et unités prosodiques**

Comme explicité dans la section [1.4](#) du chapitre, la formation de groupements d'éléments nécessite une tête, qui se distingue par sa saillance par rapport aux autres éléments au sein du groupe. Il est donc évident que l'accentuation joue un rôle fondamental dans l'émergence de ces groupements. Ce rôle de l'accentuation est étroitement lié à sa fonction démarcative (Troubetzkoy, 1949). Cette fonction marque les frontières des mots ou des groupements de mots, principalement dans des langues comme le français, où la position des accents est fixe dans une unité prosodique.

Il est crucial de noter que les accents ne sont pas disposés au hasard dans le discours. C'est ici que la notion d'organisation de métrique entre en jeu, étant chargée d'orchestrer l'emplacement des accents. La métrique vise à instaurer des régularités systématiques au sein du flux de parole (Cornulier, 1995). La phonologie métrique (Lieberman et Prince, 1977 ; Hayes, 1980, 1984 ; Hogg et McCully, 1987), place l'accentuation et l'organisation hiérarchique des énoncés au cœur de sa théorie. Elle postule que différents niveaux d'accents correspondent à différents niveaux de prééminence, et que les accents sont associés aux syllabes pour former des unités prosodiques de diverses tailles.

Ces unités prosodiques émergent en suivant deux principes métriques fondamentaux : l'alternance et la culmination (Delais-Roussarie, 2022). Le principe d'alternance exige que les éléments forts et faibles (par exemple, les syllabes accentuées et non accentuées) alternent de manière systématique, garantissant ainsi l'eurythmie dans les énoncés. Le principe de culmination souligne que les frontières des unités, quelle que soit leur taille (ou niveau), doivent être marquées par des tempos forts (tels que les accents). Il convient de noter que le nombre et la dénomination des unités prosodiques aux différents niveaux ne sont pas toujours uniformes entre les chercheurs. La section [2.2.1](#) de ce chapitre précisera les niveaux d'unités prosodiques du français étudiés dans le cadre de cette thèse.

## 2.2 Unités prosodiques du français

### 2.2.1 Application de l'accentuation en français

Le français ne possède pas d'accent lexical. Par conséquent, certains linguistes, tels que Jun et Fougeron (2000), ont avancé que le domaine élémentaire d'application de l'accent primaire (cf. [2.2.2](#)) est le syntagme accentuel (l'unité rythmique selon Di Cristo et Hirst, 1993). Ce syntagme accentuel (ex. « une jolie maison »), est constitué d'un mot lexical portant l'accent primaire et d'une ou plusieurs syllabes qui le précèdent et en dépendent (Di Cristo et Hirst, 1993). L'émergence de cette unité, le syntagme accentuel, est fortement tributaire des informations syntaxiques (Delais-Roussarie et Post, 2008). Cependant, Astésano (2017) a proposé, pour le français, le mot prosodique comme le domaine élémentaire d'application de l'accent primaire. Le mot prosodique est constitué d'un mot lexical et de ses éventuels clitiques (ex. « une maison »).

Les deux propositions concernant le domaine élémentaire d'application impliquent que l'accent primaire est associé à un groupe de mots. En effet, lorsque deux mots lexicaux se trouvent dans un même groupe, comme dans l'exemple « une jolie maison » où « jolie » et « maison » sont des mots lexicaux, le mot lexical en avant-dernier (l'adjectif « jolie ») perd son accentuation. Seule la dernière syllabe du dernier mot lexical du groupe (le nom « maison »), porte l'accent primaire (Astésano, 2017). Ainsi, dans une certaine mesure, nous pouvons considérer un mot prosodique comme un cas particulier d'un syntagme accentuel.

Le niveau supérieur du syntagme accentuel est le syntagme intonatif (Jun et Fougeron, 2000 ; égal à l'unité intonative selon Di Cristo et Hirst, 1993). Il est composé de plusieurs syntagmes

accentuels. La définition du syntagme intonatif est fortement influencée par des paramètres syntaxiques et sémantiques (Delais-Roussarie et Post, 2008). Nous n'allons pas décrire en détail tous les niveaux des unités prosodiques possibles en français. Il est essentiel de noter que notre intérêt se porte sur le rôle de l'accentuation et la hiérarchisation de l'accentuation dans le groupement des éléments lors de la production de la parole. Par conséquent, dans cette thèse, les unités telles que le syntagme accentuel (ou le mot prosodique) et le syntagme intonatif seront les principales unités prosodiques abordées. Nous exposerons donc les caractéristiques de ces deux unités dans la prochaine section ([2.2.2](#)).

### **2.2.2 Système métrique du français**

Selon Di Cristo (2003), le système métrique du français repose sur deux caractéristiques fondamentales. Premièrement, le syntagme accentuel est susceptible d'être bipolaire. En d'autres termes, chaque syntagme accentuel possède systématiquement une tête métrique à sa frontière droite, c'est-à-dire sur la dernière syllabe de chaque unité accentuelle. Cette tête métrique correspond à l'accent primaire. Toutefois, un syntagme accentuel peut également comporter un accent initial sur les premières syllabes de l'unité (Jun et Fougeron, 2000). Il convient de préciser que l'accent initial est souvent considéré comme un accent secondaire, étant donné que sa présence est soumise à diverses contraintes, telles que le nombre de syllabes dans un syntagme accentuel et la position de ce syntagme. Lorsqu'une situation de « collusion accentuelle » se produit, en violation du principe d'alternance (cf. chapitre I [2.1](#)), c'est en général l'accent final qui prédomine. Ainsi, c'est l'accent final qui joue un rôle prépondérant dans l'assurance de la fonction démarcative (cf. chapitre I [2.1](#)).

La deuxième caractéristique du système métrique du français réside dans la différence de projection en surface des deux types d'accent, à savoir l'accent initial et l'accent final. La mise en évidence de la prééminence de l'accent final peut se traduire par une variation mélodique et un allongement de la syllabe portant l'accent (Di Cristo et Hirst, 1993 ; Di Cristo, 2003 ; Jun et Fougeron, 2000). En ce qui concerne l'accent initial, les premières études n'avaient observé qu'une variation mélodique (Di Cristo, 1978 ; Rossi et Di Cristo, 1980, Rossi, 1985). Toutefois, les recherches ultérieures (Jankowski et al., 1999 ; Astésano, 2001) ont montré que la prééminence de l'accent initial pouvait également s'exprimer par le biais de la durée. La distinction entre l'accent initial et l'accent final réside dans la distribution infra-syllabique de l'allongement. Plus précisément,

l'accent initial se manifeste par un allongement de l'attaque des syllabes, tandis que pour l'accent final, l'allongement concerne davantage la rime des syllabes, c'est-à-dire le noyau et le coda.

La tête du syntagme intonatif est également alignée à droite, et elle se caractérise par un niveau de proéminence plus élevé (Di Cristo, 2003 ; Jun et Fougeron, 2000). En d'autres termes, l'accent primaire au niveau du syntagme intonatif peut se manifester par un allongement plus prononcé et un contour mélodique majeur. L'application du principe d'alternance (cf. chapitre I [2.1](#)) pourrait être moins observée dans la formation de ce constituant prosodique.

### **2.2.3 Caractéristiques de l'accentuation du français**

Il est pertinent de noter que le français a été historiquement considéré comme une langue dépourvue d'accent. Selon Di Cristo (2011), cette confusion découle de trois caractéristiques des accents en français. Tout d'abord, en français, les accents n'ont pas de fonction distinctive, c'est-à-dire que la position de l'accent au sein d'un mot ne détermine pas sa signification. De plus, par rapport aux langues germaniques (ex. l'anglais, l'allemand, le néerlandais), les distinctions entre les syllabes accentuées et non accentuées sont moins marquées. Par exemple, il a été observé que les différences de durée entre les syllabes accentuées et non accentuées étaient moins prononcées en français que dans d'autres langues (Kruckenberg et Fant, 1995 ; Astésano, 1999). Enfin, le phénomène du syncrétisme (Rossi, 1980) entre l'accentuation et l'intonation est fréquent en français. L'accent primaire est parfois caractérisé par un simple contour mélodique (Di Cristo, 2003). Par conséquent, l'accentuation joue un rôle significatif dans la modulation de l'intonation en français (Di Cristo et Hirst, 1996). Il est donc naturel de confondre ces deux phénomènes. En effet, par exemple, dans le modèle d'intonation proposé par Jun et Fougeron (2020), l'intonation exerce une influence notable dans la délimitation des unités prosodiques.

En effet, il est indiscutable que les Français produisent et identifient les accents (Astésano et al., 2004 ; Michelas et al., 2016). De plus, ils parviennent à percevoir les accents finals, indépendamment de la présence d'une frontière prosodique clairement marquée par un contour intonatif (Astésano et al., 2012). Néanmoins, la présence des accents ne favorise pas la reconnaissance des mots pour les Français (Michelas et Dufour, 2019).

## 2.3 Formalisation des phénomènes métriques

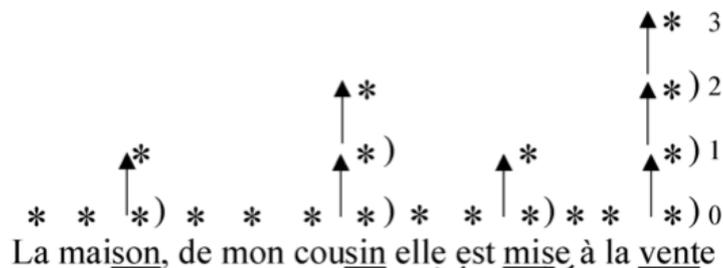
Après une revue du rôle de l'accentuation et des unités prosodiques dans la langue française, il est maintenant temps d'examiner la manière dont la métrique contribue à établir le rythme de la parole à différents niveaux d'accentuation.

Les chercheurs en linguistique ont élaboré des outils de formalisation visant à représenter les phénomènes métriques. À titre d'exemple, nous présenterons la grille métrique étendue, adaptée en français par Di Cristo (cf. Di Cristo et Hirst, 1993 ; Di Cristo, 2016). La grille métrique trouve ses origines dans le travail de Prince (1983) et permet de rendre compte des relations de proéminence encodées par le système phonologique. La version étendue de cette grille, proposée par Halle et Vergnaud (1987), permet également de faire référence à la structuration des unités à différents niveaux de la prosodie. À titre d'illustration, la Figure 4 présente un exemple de grille métrique étendue pour l'analyse de la phrase « La maison de mon cousin, elle est mise à la vente », tirée de Di Cristo (2016, p. 61). Dans une grille métrique, le niveau de proéminence de chaque syllabe est représenté par le nombre d'étoiles « \* » (c'est-à-dire la hauteur des colonnes). Par exemple, la dernière syllabe, « vente », de la phrase est marquée par le plus grand nombre d'étoiles « \* », ce qui signifie qu'elle est métriquement la plus proéminente de la phrase. Dans cette version de la grille métrique, le niveau de proéminence, marqué par les symboles « \* », reflète principalement la hauteur de la voix, car la hauteur de voix est un paramètre couramment utilisé en français pour noter les accents. Dans le contexte de la langue française, des parenthèses droites sont employées pour indiquer que la syllabe spécifiée, ainsi que les syllabes qui la précèdent, appartiennent à la même unité prosodique. Par exemple, la parenthèse à droite de la syllabe « son » signifie que les trois syllabes « La maison » font partie de la même unité prosodique.

La ligne 1 de la grille projette ainsi les têtes métriques des domaines délimités dans la ligne 0, tandis que la ligne 2 projette les têtes métriques des domaines délimités dans la ligne 1, et ainsi de suite. Dans l'exemple donné, les domaines délimités dans la ligne 0 (« la maison », « de mon cousin », « elle est mise », et « à la vente ») peuvent correspondre soit aux mots prosodiques d'Astésano (2017), soit aux syntagmes accentuels tels que définis par Jun et Fougeron (2000), ou encore aux unités rythmiques selon Di Cristo et Hirst (1993) (cf. chapitre I 2.2.1). Ainsi, les domaines délimités dans la ligne 1 (« la maison de mon cousin » et « elle est mise à la vente ») correspondent aux syntagmes intonatifs. En fin de compte, le niveau supérieur, représenté dans la ligne 2, englobe l'ensemble de l'énoncé.

Figure 4 Illustration de la grille métrique étendue extrait de Di Cristo (2016)

Le nombre d'étoiles (\*) figurant dans chaque colonne reflète le niveau relatif de proéminence de chaque syllabe. Les parenthèses droites servent à délimiter les domaines métriques à chaque niveau.



Les avantages de la formalisation de la grille métrique étendue sont manifestes. En premier lieu, elle permet de mettre en exergue les deux principes métriques fondamentaux qui sous-tendent la formation des unités prosodiques : l'alternance et la culmination (cf. chapitre I [2.1](#)). En conséquence, la grille métrique offre, en théorie, une représentation du groupement rythmique, c'est-à-dire l'organisation des syllabes en unités de diverses tailles (ex. le groupement des syllabes « la maison » en syntagme accentuel). De surcroît, bien que, dans le contexte du français, le niveau de proéminence dans la grille métrique corresponde souvent au niveau de proéminence tonale, il est toutefois aisé d'intégrer d'autres paramètres prosodiques, tels que la durée et l'intensité, dans l'analyse de la proéminence dans la grille. Cependant, dans cette version de la grille métrique, les accents initiaux (cf. chapitre I [2.2.2](#)) ne sont pas pris en considération.

## 2.4 Rythme dérivant de la prosodie tout en s'en distinguant

La grille métrique, présentée dans la section [2.3](#), établit un lien entre la structure prosodique et le groupement rythmique de la parole. Les règles prosodiques déterminent le niveau relatif de proéminence des syllabes au sein d'un énoncé, ce qui permet de grouper ces syllabes en unités de diverses tailles. La prosodie, par conséquent, joue un rôle crucial dans la formation du groupement rythmique.

Il est néanmoins impératif de souligner que, même si les contraintes métriques occupent une place centrale dans la génération du rythme, l'objectif sous-jacent à l'organisation du rythme dépasse largement la simple structuration de l'accentuation. Du point de vue de la construction rythmique, il convient de noter que toutes les proéminences observées en surface ne sont pas nécessairement liées à l'accentuation prosodique. À titre d'exemple, il arrive que les locuteurs

mettent en évidence une syllabe qui, selon la métrique, ne doit pas être accentuée, dans le but de mettre en relief des informations spécifiques transmises par cette syllabe lors de la communication (Delais-Roussarie, 2022). Par ailleurs, comme abordé dans la section [1](#) de ce chapitre, le rythme émerge de la coordination des mouvements de multiples articulateurs, soumis à des contraintes biomécaniques, neurophysiologiques et cognitives (MacNeilage, 1998 ; Davis et MacNeilage, 2004 ; Di Cristo, 2003 ; Assaneo et Poeppel, 2018).

Ainsi, bien que la prosodie ait un impact sur la modulation du rythme, il est important de reconnaître que le rythme est un phénomène dynamique au sein de la communication.

## 2.5 Mesurer des régularités rythmiques

Les recherches portant sur le rythme linguistique ont longtemps été consacrées à l'analyse de l'isochronie des intervalles entre les accents (ex. Pike, 1945). Ces études ont avancé des postulats prévoyant une distinction fondamentale entre l'isochronie accentuelle et l'isochronie syllabique au sein des langues. Selon cette distinction, les langues qui adoptent l'isochronie accentuelle se caractérisent par des intervalles réguliers entre les accents, tandis que les langues à l'isochronie syllabique présentent des intervalles réguliers entre les syllabes. À titre d'exemple, l'anglais et l'allemand sont classés parmi les langues à l'isochronie accentuelle, tandis que le français, quant à lui, est catégorisé en tant que langue à l'isochronie syllabique.

Toutefois, des recherches ultérieures n'ont pas corroboré cette prédiction (cf. Arvaniti, 2012 pour une revue de la littérature). C'est à partir des travaux de Dauer (1983, 1987), que l'orientation des études sur le rythme a évolué de la simple mesure des durées vers la mise en avant de la prééminence des accents. Selon la définition de Dauer, le rythme se manifeste par le groupement des éléments en unités plus larges, concept en adéquation avec la notion de groupement rythmique proposée dans la présente thèse. Selon Dauer, du moins dans certaines langues, les accents sont utilisés comme points d'ancrage perceptif pour former des groupements syllabiques. De plus, plusieurs paramètres, tels que la réduction des voyelles, la complexité de la structure syllabique, et la réalisation phonétique de la prééminence prosodique, peuvent être mobilisés pour faciliter ce groupement. Par conséquent, il est établi que le groupement rythmique est une caractéristique présente dans toutes les langues. Les langues parlées par les êtres humains peuvent être organisées le long d'un continuum en ce qui concerne le degré de groupement rythmique.

Par la suite, une nouvelle approche visant à mesurer le rythme en quantifiant la variabilité des intervalles entre les voyelles et les consonnes a été développée (Ramus et al., 1999 ; Low et al., 2000 ; Grabe et Low, 2002). Cette perspective de mesure du rythme recentre l'attention sur le paramètre de la durée. Les métriques les plus couramment utilisées sont le pourcentage des intervalles vocaliques dans les énoncés (%V) et l'écart-type des durées consonantiques ( $\Delta C$ ) introduits par Ramus et al. (1999), ainsi que l'indice des comparaisons en paire des intervalles vocaliques et intervocaliques successifs (nPVI) développé par Low et al (2000) (cf. également Grabe et Low, 2002). Plusieurs variantes de ces métriques ont également été proposées (ex. Dellwo, 2006 ; Bertinetto et Bertini, 2008). Cependant, ces métriques rythmiques ont suscité des controverses quant à leur capacité à classer de manière fiable les langues du monde. Ces désaccords résultent, entre autres, de la grande variabilité interlocuteur, du choix des corpus étudiés, et des méthodes de collecte de données (Arvaniti, 2009).

En effet, ces métriques rythmiques (Ramus et al., 1999 ; Low et al., 2000 ; Grabe et Low, 2002) ont été élaborées en se fondant sur les paramètres, tels que la durée syllabique et la réduction des voyelles, proposés par Dauer (1983, 1987) pour structurer les langues le long d'un continuum rythmique. Cependant, il est essentiel de noter que ces métriques ne mettent en lumière que l'aspect de la variation des durées, sans prendre en compte le fait que les paramètres avancés par Dauer visent principalement à rendre compte du degré de facilité à définir des proéminences dans une langue (Arvaniti, 2009). Cela implique également que les métriques rythmiques classiques n'ont pas été en mesure de rendre compte de la hiérarchisation du rythme et du groupement des éléments en unités de différentes tailles.

L'application des métriques rythmiques basées sur la durée peut être entravée par une source de complexité importante, résultant du fait que les locuteurs ont tendance à réguler la durée relative des éléments linguistiques plutôt que leur durée intrinsèque. À titre d'exemple, l'étude de Hermes et al. (2020) a mis en évidence que la durée des consonnes géminées (deux consonnes identiques successives, telles que /tt/ dans le mot japonais « kitte » [« timbre »]) et des consonnes simples (par exemple /t/ dans le mot japonais « kite » [« personne qui vient »]) se distinguait en fonction de la proportion de la durée du mot, indépendamment du débit de la parole.

Néanmoins, la question fondamentale se pose quant à la normalisation des mesures absolues. En raison de l'influence de divers facteurs, tels que le débit de la parole et le style du discours, sur la durée des unités linguistiques, il subsiste une incertitude quant à la référence à adopter pour la

normalisation de ces mesures. Hermes et al. (2020) ont opté pour la normalisation de la durée des consonnes par rapport à la durée du mot, tandis que Grabe et Low (2002) ont choisi de normaliser l'intervalle vocalique par rapport à l'intervalle vocalique précédent. Il est quasiment impossible de tenir compte de tous les facteurs qui influencent la durée des éléments linguistiques pour évaluer la référence à utiliser dans la normalisation de la durée.

Par conséquent, il convient de souligner que la mesure de la durée relative ne peut être utilisée que dans des cas contrôlés, où un ou peu de facteurs sont systématiquement variés entre les conditions expérimentales, afin de caractériser les régularités rythmiques.

### **3. Prosodie en tant que mécanisme stabilisant du contrôle moteur**

Dans la section [2](#) de ce chapitre, nous avons examiné la manière dont la structure prosodique influe sur la structure rythmique, tout en notant que les patrons rythmiques de surface ne se conforment pas toujours strictement aux principes métriques. La prosodie facilite l'émergence du groupement rythmique (c'est-à-dire l'organisation des éléments de la production de la parole, tels que les syllabes) à différentes échelles temporelles. Il est à présent opportun de revisiter la fonction du groupement, dont nous avons discuté dans la section 1 de ce chapitre, dans le but d'évaluer comment la prosodie contribue à l'accomplissement de cette fonction de groupement.

#### **3.1 Rythme en tant que système coordinatif**

Conformément à ce qui a été abordé dans la section [1.2](#) du chapitre précédent, le groupement des éléments de la production de la parole (ex. les gestes articulatoires), sert à les intégrer au sein d'une structure temporelle commune (c'est-à-dire le rythme). Cette intégration vise à minimiser le niveau de contrôle requis, réduisant ainsi la complexité du système de contrôle. Il est également pertinent de noter que le groupement s'organise de manière hiérarchique (cf. Figure [3](#)). Ces observations impliquent non seulement une coordination des éléments au même niveau, mais également une étroite interaction entre les composantes de différents niveaux. En s'appuyant sur

cette perspective, de nouvelles approches d'études rythmiques ont été proposées (ex. Cummins et Port, 1998). Cette approche considère le rythme comme un système coordinatif, au sein duquel les éléments à différentes échelles temporelles se coordonnent et sont contraintes dans leur relation temporelle. Ainsi, cette perspective d'étude rythmique ne se concentre pas uniquement sur des éléments spécifiques au sein des énoncés, mais plutôt sur les interactions entre ces éléments.

La prosodie engendre la création de plusieurs niveaux de groupement et influence la coordination des éléments au sein de ces groupements. Comme exposé dans la section [2](#), le système prosodique délimite les frontières du groupement (c'est-à-dire l'unité prosodique), à travers le processus d'accentuation, qui confère une proéminence métrique à certaines syllabes au sein des énoncés. En effet, l'accentuation peut même exercer une influence directe sur les mouvements des articulateurs. Par exemple, les mouvements articulatoires nécessaires pour produire des syllabes accentuées peuvent être plus amples (Smith et al., 2019 ; Meynadier et al., 1998), et une moindre coordination peut être observée lors de la production de syllabes accentuées (De Jong et al., 1993).

Par conséquent, il est possible de considérer la prosodie comme un mécanisme visant à stabiliser le contrôle moteur lors de la production de la parole, en favorisant l'émergence du groupement rythmique. Parallèlement, elle régule la coordination des éléments au sein du même niveau de groupement ainsi que la coordination des éléments de niveaux différents.

## **3.2 Coordination entre les composantes rythmiques à différentes échelles temporelles**

Comme exposé dans la section [2.4](#) de ce chapitre, le rythme de surface se caractérise par sa dynamique, car il est ajusté en temps réel par les locuteurs au cours de la production de la parole afin de l'adapter à la situation actuelle. En conséquence, le système coordinatif mis en place est spécifique à chaque tâche de production de la parole en fonction du contexte actuel.

### **3.2.1 Composante syllabique et composante supra-syllabique**

Au sein de cette thèse, notre attention est concentrée sur la coordination entre deux niveaux de composantes spécifiques : le niveau de la composante syllabique et le niveau de la composante groupant deux ou plusieurs syllabes. Nous ne faisons pas immédiatement correspondre ces deux niveaux de composantes aux unités prosodiques dont nous avons discuté dans la section [2](#). Notre

raisonnement découle de la considération du groupement rythmique comme un processus qui émerge en surface pendant la production de la parole.

La composante syllabique se forme à partir de la production des syllabes (cf. chapitre I [1.3](#)). La prosodie intervient en rendant certaines syllabes plus proéminentes que d'autres qui les entourent. Ce processus contribue à l'apparition d'une composante de niveau supérieur à celui des syllabes. Cette perspective sur l'émergence du groupement rythmique est cohérente avec l'idée des analyses des unités métriques de la grille métrique proposées dans le cadre de la phonologie métrique (Halle et Vergnaud, 1987), comme évoqué dans la section [2.3](#) de ce chapitre. Cependant, il est important de noter que toutes les syllabes proéminentes observées en surface ne sont pas nécessairement associées à une accentuation prosodique. Par conséquent, nous utilisons le terme « composante supra-syllabique » pour désigner la composante située au-dessus de la composante syllabique, afin de mieux rendre compte du fait que son émergence est liée aux syllabes proéminentes observées en surface.

Dans une certaine mesure, cette perspective d'analyse du groupement rythmique dans la parole rejoint l'hypothèse de Dauer (1983, 1987) selon laquelle le groupement rythmique existe dans toutes les langues du monde, car les syllabes ne peuvent pas avoir le même niveau de proéminence.

### **3.2.2 Coordination entre la composante syllabique et la composante supra-syllabique**

La composante syllabique est imbriquée au sein de la composante supra-syllabique. Au cours de la production de la parole, ces deux composantes doivent être interconnectées dans un état stable propre à la langue parlée et aux conditions de production de la parole (Cummins et Port, 1998). Par exemple, une étude menée par Lancia et al. (2019) a révélé une variabilité du degré de coordination (défini comme la stabilité de la relation temporelle entre deux signaux continus dans le temps) entre les modulations syllabiques et supra-syllabiques en fonction des langues. Ces modulations d'amplitude correspondent aux enveloppes d'amplitude, qui représentent les oscillations indiquant les macro-fluctuations d'amplitude respectivement à l'échelle temporelle syllabique et à l'échelle temporelle supra-syllabique (pour plus de détails, cf. chapitre III [3.2](#) dédié à la description méthodologique). L'étude de Lancia et al. (2019) a constaté que le degré de coordination entre la composante syllabique et la composante supra-syllabique était généralement plus élevé dans les énoncés des langues dites à « rythme accentuel » telles que l'anglais et l'allemand, que dans ceux des langues dites à « rythme syllabique » telles que le français et l'italien. En d'autres

termes, un système coordinatif spécifique à la structure prosodique de chaque langue doit être généré pendant la production de la parole.

Un degré élevé de coordination entre la composante syllabique et la composante supra-syllabique a été établi en concordance avec le degré de groupement syllabique. À titre d'exemple, d'une part, les travaux de Leong et al. (2014) ont révélé que, dans le contexte de la parole adressée aux enfants, une synchronisation temporelle plus prononcée se manifestait entre les modulations d'amplitude du rythme syllabique (2,5 - 12 Hz) et celles du rythme supra-syllabique (0,9 - 2,5 Hz) des signaux d'amplitude de la parole. D'autre part, la parole adressée aux enfants se caractérise par un degré de groupement plus élevé entre les pics d'amplitude à différentes échelles temporelles (c'est-à-dire à l'échelle syllabique et à l'échelle supra-syllabique, Falk et Kello, 2017).

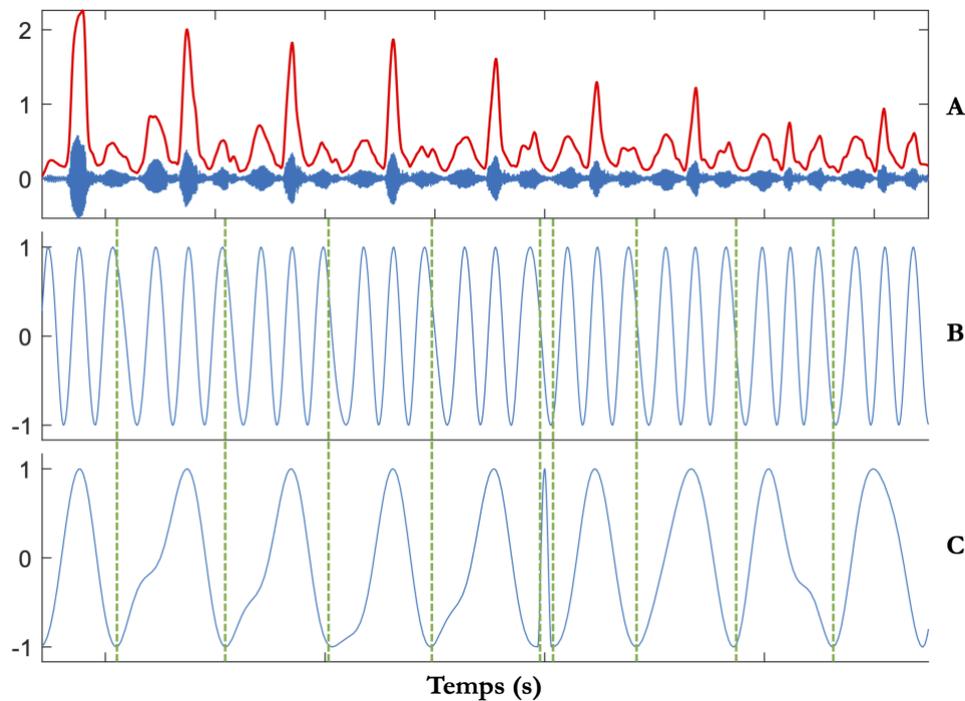
Notre étude récente (Lancia et al., 2023) a également mis en évidence cette relation entre la coordination des composantes rythmiques et le groupement rythmique. La Figure 5, extraite de notre recherche, illustre le signal acoustique d'un extrait de la répétition de la séquence de syllabes /badego/ (panneau A), ainsi que les composantes syllabique (l'oscillation de l'énergie liée à la production des syllabes - panneau B) et supra-syllabique (l'oscillation de l'énergie liée à la production des syllabes plus proéminentes - panneau C). Ces deux composantes ont été extraites en filtrant le signal d'amplitude et en procédant au traitement de chaque cycle pour obtenir des oscillations comprises dans la plage de -1 à 1 (voir chapitre III 3.2 pour plus de détails). Il est observé que la présence des syllabes proéminentes tend à grouper les syllabes par groupes de trois, ce qui se traduit par une relation temporelle régulière entre l'oscillation syllabique (B) et l'oscillation supra-syllabique (C). Plus précisément, les cycles lents (supra-syllabiques) modulent de manière stable les cycles rapides (syllabiques) dans le temps, avec environ un cycle lent correspondant à trois cycles rapides. Les cycles rapides modulés par un même cycle lent présentent donc un degré d'interdépendance plus élevé.

Du point de vue perceptif, la parole adressée aux enfants est généralement perçue comme étant plus rythmique. Une parole rythmique est fréquemment associée à un groupement plus fort et régulier des éléments consécutifs. En effet, l'étude de Leong et Goswami (2014) a mis en évidence une relation entre le degré de coordination des composantes rythmiques et la rythmicité perceptive. En d'autres termes, du point de vue perceptif, le degré de couplage entre les composantes rythmiques est également corrélé au degré de groupement perceptif entre les éléments consécutifs.

L'exemple de la parole adressée aux enfants illustre comment le groupement rythmique peut s'adapter en fonction de l'interlocuteur, tel que l'enfant, ainsi que de l'intention de la communication, telle que capter l'attention de l'enfant. Cette adaptation du groupement rythmique pourrait se traduire par l'ajustement du degré de coordination entre les composantes rythmiques à différentes échelles temporelles, permettant ainsi d'atteindre un état stable de coordination entre les composantes en fonction de la situation de production de la parole.

Figure 5 Illustration de la corrélation entre le groupement syllabique et la relation temporelle entre la composante syllabique et la composante supra-syllabique

A : le signal acoustique d'un extrait de la répétition d'une séquence de syllabes /badego/ ; B : la composante syllabique correspondant à l'oscillation de l'énergie au rythme de la production des syllabes ; C : la composante supra-syllabique correspondant à l'oscillation de l'énergie au rythme de la production des syllabes proéminentes



## 4. Résumé du chapitre

L'une des sources importantes de complexité dans la production de la parole réside dans la nécessité d'orchestrer plusieurs articulateurs dans le contrôle moteur. Pour atténuer cette complexité, les articulateurs coordonnent leurs mouvements en vue d'atteindre un objectif linguistique commun. Les mouvements des articulateurs coordonnés présentent une forte corrélation, générant ainsi un rythme coordinatif (Poeppel et Assaneo, 2020). Ce rythme

coordonatif contribue à réduire les dimensions de contrôle au cours de la production de la parole, simplifiant ainsi le contrôle moteur. De plus, cette coordination permet au système moteur de réagir rapidement aux perturbations survenues pendant la production de la parole (Kelso et al., 1984).

Le rythme coordinatif généré par la coordination des mouvements des articulateurs correspond généralement à la production des syllabes (cf. Figure 1). Le rythme syllabique est complexe, non seulement en raison de son origine dans la coordination de plusieurs articulateurs, mais aussi du fait qu'il est soumis à diverses contraintes d'ordre biomécanique, neurophysiologique et cognitif (MacNeilage, 1998 ; Davis et MacNeilage, 2004 ; Di Cristo, 2003 ; Assaneo et Poeppel, 2018).

De plus, le rythme syllabique est également modulé par des facteurs prosodiques, qui confèrent une plus grande saillance à certaines syllabes au sein d'un énoncé à travers le processus d'accentuation. Ces syllabes saillantes divisent le flux de la parole en unités et servent de têtes à ces unités. L'accentuation est organisée hiérarchiquement, avec différents niveaux délimitant des tailles d'unités distinctes. En français, l'unité minimale qui porte un accent peut être le mot prosodique (Astésano, 2017), ou le syntagme accentuel (Jun et Fougeron, 2000). Deux ou plusieurs syntagmes accentuels peuvent être groupés pour former un syntagme intonatif (Jun et Fougeron, 2000). Modulé par la prosodie, le groupement rythmique (le groupement des syllabes) est également organisé hiérarchiquement, en relation avec l'organisation de plusieurs niveaux de proéminence, qui peut être représenté par une grille métrique (cf. Figure 4).

Cependant, la structure rythmique de surface observée peut différer de la structure de proéminences (liée à l'accentuation). D'une part, cela est dû au fait que le rythme est également soumis à plusieurs autres contraintes, comme nous l'avons mentionné précédemment. D'autre part, le rythme est flexible et dynamique, s'adaptant à la situation de production de la parole. Il est très difficile de mesurer les régularités rythmiques uniquement à l'aide des métriques basées sur la durée, lesquelles ne prennent pas en compte la hiérarchisation du rythme. En plus, même lorsque nous mesurons une durée relative, il reste compliqué de déterminer la référence à utiliser pour calculer cette durée relative.

Dans ce contexte, nous proposons d'adopter une nouvelle perspective pour étudier les patrons rythmiques, dans laquelle le rythme est considéré comme un mécanisme coordinatif. Dans ce mécanisme, la composante syllabique et la composante supra-syllabique (cf. Figure 5) se coordonnent et sont contraintes dans leur relation temporelle. Le degré de coordination entre ces deux composantes est cohérent avec le degré de groupement syllabique (Leong et al., 2014 ; Falk

et Kello, 2017, Lancia et al., 2023), et il a été constaté comme différent d'une langue à l'autre (Lancia et al., 2019). L'ajustement des patrons rythmiques en fonction de la situation de communication consiste donc à ajuster le degré de coordination entre ces deux composantes.

Il est important de noter que pour garantir la stabilité et la flexibilité du système de production de la parole, un système d'intégration des informations sensorielles dans le contrôle moteur est sollicité. Dans le prochain chapitre, nous explorerons ce système sensorimoteur et examinerons les facteurs qui influencent les processus de contrôle sensorimoteur.



# CHAPITRE II : FLEXIBILITE DE LA PRODUCTION DE LA PAROLE

## Sommaire du chapitre

<b>1. Contrôle sensorimoteur pendant la production de la parole</b> .....	<b>30</b>
1.1 Processus du contrôle sensorimoteur dans la production de la parole .....	30
1.1.1 Intégration sensorimotrice .....	30
1.1.2 Intégration des retours sensoriels et du mécanisme prédictif.....	31
1.2 Retour auditif suivi pendant la production de la parole .....	33
1.2.1 Dimensions du retour auditif suivies .....	33
1.2.2 Intégration du retour auditif et somatosensoriel .....	35
<b>2. Contrôle sensorimoteur de la production de la parole interagissant avec la complexité linguistique</b> .....	<b>36</b>
2.1 Contrôle sensorimoteur modulé par l'expérience linguistique.....	36
2.2 Rôle de l'accentuation dans le contrôle sensorimoteur .....	37
2.3 Interaction entre organisation rythmique et contrôle sensorimoteur en production de la parole : perspectives.....	38
<b>3. Rôle du sens d'agentivité dans le contrôle sensorimoteur pendant la production de la parole</b> .....	<b>39</b>
3.1 Effet des entrées externes sur la production de la parole.....	39
3.1.1 Convergence phonétique.....	39
3.1.2 Cibles sensorimotrices .....	40
3.2 Génération du sens d'agentivité implicite .....	41
3.2.1 Définition du sens d'agentivité .....	41
3.2.2 Lien entre le sens d'agentivité implicite et le contrôle sensorimoteur .....	42
3.2.3 Génération du sens d'agentivité implicite pendant la production de la parole.....	44
3.3 Flexibilité du sens d'agentivité implicite .....	48
3.3.1 Compétition entre les cibles externes et internes.....	48
3.3.2 Altération du sens d'agentivité implicite .....	49
3.4 Sens d'agentivité implicite modulant le contrôle sensorimoteur de la parole.....	50
<b>4. Résumé du chapitre</b> .....	<b>51</b>

En dépit de sa complexité intrinsèque, le système de la production de la parole démontre une remarquable flexibilité face à diverses conditions. Cette flexibilité est illustrée par notre capacité à ajuster notre volume vocal en réponse à un environnement bruyant et à maintenir la parole en mangeant (même si cela peut être considéré comme impoli envers notre interlocuteur). Cette flexibilité est, en partie, favorisée par l'émergence du rythme, comme abordé dans le chapitre I, lequel simplifie le contrôle moteur. Toutefois, il est essentiel de noter que cette flexibilité repose

largement sur le système de contrôle sensorimoteur, sujet que nous approfondirons dans la section [1](#) de ce chapitre. Étant donné l'importance de la prosodie et du rythme dans le contrôle moteur, la section [2](#) examinera de près les interactions entre les processus sensorimoteurs et les facteurs prosodiques. Finalement, dans la section [3](#), nous considérerons le fait que les mécanismes du contrôle sensorimoteur peuvent différer lorsqu'ils sont sollicités dans un contexte d'interaction verbale. Cela nous amènera à explorer la notion du sens d'agentivité, qui en parole concerne notre capacité à distinguer notre propre voix de celle d'autrui. Nous analyserons donc la manière dont le sens d'agentivité se génère dans les processus du contrôle sensorimoteur, ainsi que son rôle dans la modulation de ces processus.

## **1. Contrôle sensorimoteur pendant la production de la parole**

Afin de garantir des mouvements précis et fluides, le système moteur intègre les informations provenant des récepteurs sensoriels telles que les informations visuelles, auditives et somatosensorielles, permettant ainsi d'ajuster en temps réel les mouvements. Dans cette section, nous aborderons tout d'abord le mécanisme du contrôle sensorimoteur pendant la production de la parole. Nous nous concentrerons ensuite sur le rôle essentiel du retour sensoriel, en particulier le retour auditif, dans le contrôle sensorimoteur lors de la production de la parole.

### **1.1 Processus du contrôle sensorimoteur dans la production de la parole**

#### **1.1.1 Intégration sensorimotrice**

Au quotidien, notre capacité à générer la parole dans un environnement en constante évolution repose fondamentalement sur l'intégration sensorimotrice. Plus précisément, le système de production de la parole procède à une évaluation continue de l'état courant des mouvements des articulateurs en comparant les retours sensoriels, c'est-à-dire la perception des informations sensorielles résultant de nos propres actions, avec les conséquences sensorielles souhaitées (Parrell et al., 2019). Cette évaluation ininterrompue de son état permet ainsi au système de production de

la parole d'ajuster en temps réel son comportement, garantissant la stabilité de la production malgré les perturbations environnementales.

Pendant la production de la parole, le retour auditif et le retour somatosensoriel jouent un rôle prépondérant. Le retour auditif fait référence à la perception des signaux sonores résultant de nos propres articulations. Quant au retour somatosensoriel, dans le contexte des études sur la parole, il renvoie à la configuration et à la dynamique du conduit vocal et est généralement corrélé au retour auditif. Ces deux types de retour fournissent des informations complémentaires qui contribuent au contrôle de la production de la parole.

Bien que les modalités d'intégration des informations sensorielles puissent varier ou présenter des nuances au sein des différents modèles de la production de la parole (ex. Guenther et al., 2006 ; Saltzman et Munhall, 1989 ; Houde et Nagarajan, 2011 ; Parrell et al., 2019b ; Kröger et al., 2009 ; Perrier et al., 2006 ; Patri et al., 2015), il convient de noter que, dans la majorité de ces modèles, les retours sensoriels jouent un rôle crucial dans la régulation de leur modèle.

### ***1.1.2 Intégration des retours sensoriels et du mécanisme prédictif***

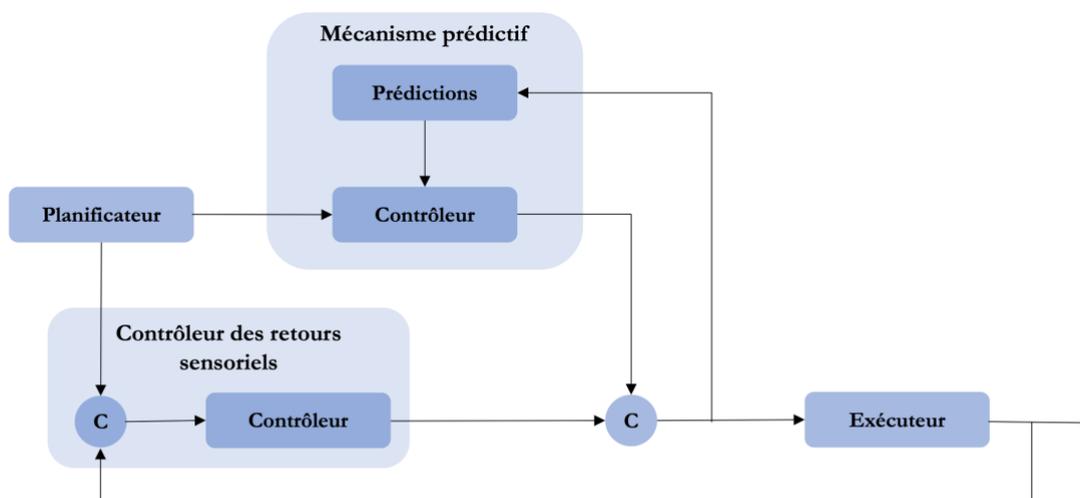
Il convient de noter que les retours sensoriels présentent des caractéristiques de lenteur et de sensibilité aux bruits. Une dépendance totale aux retours sensoriels pour estimer l'état actuel du système de production de la parole peut entraîner des inexactitudes dans la planification des mouvements (Parrell et al., 2019). Par conséquent, certains modèles ont proposé l'intégration d'un mécanisme prédictif. Ce mécanisme est en mesure de prédire les conséquences sensorielles des mouvements des articulateurs, ce qui lui confère un avantage en termes de rapidité. De plus, ce mécanisme prédictif peut fournir des informations pour l'estimation de l'état (Wolpert et al., 1995 ; Miall et Wolpert, 1996). Ainsi, dans ces modèles, les retours sensoriels et le mécanisme prédictif interviennent de manière complémentaire.

La Figure 6 illustre l'une des modalités d'intégration des retours sensoriels et du mécanisme prédictif au sein du système de production de la parole, adaptée de Parrell et al. (2019). Dans ce schéma, le « Planificateur » est responsable de la planification des objectifs linguistiques, qui sont ensuite transmis au « Contrôleur » pour les convertir en commandes motrices concrètes. Ces commandes motrices sont ensuite acheminées vers l'« Exécuteur » pour être transformées en mouvements réels des articulateurs. Au cours de ces processus, un « Contrôleur des retours sensoriels » ainsi qu'un « Mécanisme prédictif » contribuent à l'estimation en temps réel de l'état du

système, permettant ainsi une adaptation aux changements environnementaux et aux perturbations survenues. Au sein du « Contrôleur des retours sensoriels », les objectifs linguistiques planifiés sont comparés aux sorties de l'« Exécuteur » (c'est-à-dire les retours sensoriels) dans le « Comparateur » (représenté par « C » dans le cercle en bleu foncé). Les planifications sont ajustées avant d'être transmises au « Contrôleur » pour générer les commandes motrices. De l'autre côté, dans le « Mécanisme prédictif », un « Contrôleur » indépendant du « Contrôleur » dans le « Contrôleur des retours sensoriels » prédit l'état futur du système en fonction des commandes motrices actuelles. Ces prédictions sont comparées aux commandes motrices générées par le « Contrôleur » dans le « Contrôleur des retours sensoriels » dans un autre « Comparateur » (représenté par « C »), permettant ainsi d'ajuster les commandes motrices avant de les envoyer à l'« Exécuteur ». Il est à noter que les prédictions du « Mécanisme prédictif » peuvent ne pas être parfaitement exactes, mais elles contribuent à un contrôle plus rapide du système, en particulier lorsque les retours sensoriels sont retardés.

Figure 6 Schéma illustrant une modalité possible de l'intégration d'un mécanisme prédictif et des retours sensoriels dans le système de production de la parole (adapté de Parrell et al., 2019)

Le « Planificateur » est chargé de la planification des objectifs linguistiques, tandis que le « Contrôleur » au sein du « Contrôleur des retours sensoriels » convertit ces objectifs linguistiques planifiés en commandes motrices. Les ajustements des planifications sont effectués par le biais du « Comparateur » (noté « C »), qui compare les planifications du « Planificateur » aux retours sensoriels provenant de l'« Exécuteur » avant de les transmettre au « Contrôleur ». De l'autre côté, le « Contrôleur » du « Mécanisme prédictif » prédit l'état futur du système en se basant sur les commandes motrices actuelles. Ces prédictions sont ensuite confrontées aux commandes motrices générées par le « Contrôleur » au sein du « Contrôleur des retours sensoriels », et ce processus d'ajustement a lieu dans un autre « Comparateur » (également noté « C »). Enfin, ces commandes motrices ajustées sont transmises à l'« Exécuteur » pour être transformées en mouvements des articulateurs.



Il convient de noter que la Figure 6 illustre uniquement l'une des modalités d'intégration du mécanisme prédictif et des retours sensoriels. Tous les modèles de production de la parole

n'intègrent pas ces deux paramètres de la même manière. De plus, les deux types de retours sensoriels, à savoir l'auditif et le somatosensoriel, ne jouent pas le même rôle dans tous les modèles. Par exemple, dans le modèle DIVA (« *Directions into Velocities of Articulators* », Guenther et al., 2006 ; Tourville et Guenther, 2011), le retour auditif et le retour somatosensoriel sont équilibrés dans leur contribution au système. En revanche, dans le modèle TD (« *Task Dynamics* », Saltzman et Munhall, 1989), c'est le retour somatosensoriel qui occupe une place prédominante dans le processus.

## 1.2 Retour auditif suivi pendant la production de la parole

### 1.2.1 Dimensions du retour auditif suivies

Dans le cadre de cette thèse, notre attention sera principalement portée sur le rôle du retour auditif dans les processus du contrôle sensorimoteur. Le retour auditif englobe la perception de plusieurs dimensions de la parole, par exemple, la hauteur du son (c'est-à-dire la fréquence fondamentale), l'amplitude (c'est-à-dire l'intensité) et les formants.

#### Dimensions acoustiques

Les recherches portant sur la manipulation du retour auditif nous permettent de mieux comprendre comment il est suivi et comment il influence le contrôle moteur de la parole. Dans ce type d'études, les expérimentateurs enregistrent, altèrent et repassent en temps réel la voix des locuteurs pendant qu'ils produisent des sons de parole. Lorsqu'une dimension du retour auditif est altérée de manière imprévisible, la plupart des locuteurs ont tendance à compenser immédiatement leur production dans la direction opposée de l'altération. Par exemple, Burnett, Freedland, Larson et Hain (1998) ont décalé positivement ou négativement la  $f_0$  du retour auditif des participants de 100 cents pendant qu'ils produisaient la voyelle /a/. La plupart des participants ont eu tendance à adapter leur production dans la direction opposée au décalage de la  $f_0$  du retour auditif. La réaction de compensation des locuteurs au décalage imprévisible de la  $f_0$  de leur retour auditif est donc la conséquence d'une comparaison entre la  $f_0$  prédite et la  $f_0$  perçue (c'est-à-dire la  $f_0$  du retour auditif), qui engendre les erreurs de prédiction. Les commandes motrices compensatoires sont engendrées afin de réduire les erreurs (cf. [1.1](#)).

Les réponses compensatoires des locuteurs face à des altérations du retour auditif ont été étudiées dans différents contextes. Par exemple, ces phénomènes de compensation ont été

observés lorsque les formants ont été perturbés (Houde et Jordan, 1998, 2002 ; cf. également Caudrelier et Rochet-Capellan, 2019 pour une revue de la littérature), lorsque l'intensité de la parole a été altérée (Larson et al., 2007) ou que les caractéristiques spectrales des fricatives ont été modifiées (Shiller et al., 2009 ; Casserly, 2011).

### Dimensions temporelles

Des recherches ont mis en évidence que le système sensorimoteur des locuteurs suivait de manière attentive la durée des unités linguistiques qu'ils produisaient. Dans une étude récente menée par Karlin, Naber et Parrell (2021), lorsque les participants ont produit des syntagmes anglais tels que « *a capper* », « *a gapper* », « *a sapper* » et « *a zapper* », les expérimentateurs ont allongé la durée des consonnes initiales (c'est-à-dire la durée du délai d'établissement du voisement des consonnes /k, g/ et la durée des fricatives /s, z/) et réduit la durée de la voyelle /æ/ qui suivait. En conséquence, les participants ont allongé la voyelle et réduit la durée relative des consonnes dans les syllabes. Bien que la réduction de la durée relative des consonnes puisse simplement résulter de l'allongement de la voyelle, cette étude suggère que les locuteurs suivent de près les caractéristiques temporelles de leurs énoncés, tout comme les caractéristiques spectrales.

La congruence temporelle entre l'articulation et le retour auditif est également cruciale pour assurer la stabilité de la production de la parole, comme en témoignent les réponses de locuteurs au retour auditif retardé (Kalveram et Jäncke, 1989 ; Stuart et Kalinowski, 2015 ; Yates, 1963). Lorsque le retour auditif est retardé, les locuteurs entendent leur propre voix plus tard que prévu pendant la production de la parole, ce qui peut les amener à ralentir leur débit de parole (c'est-à-dire à allonger leurs syllabes) et peut entraîner une augmentation des erreurs de production et des disfluences (comme les comportements proches du bégaiement) avec un retard plus important (Black, 1951). Les locuteurs allongent la durée de leurs syllabes dans le but de minimiser le décalage perçu entre les mouvements articulatoires et leurs conséquences acoustiques. De plus, l'incongruence temporelle entre l'articulation et le retour auditif peut affecter la façon dont les locuteurs contrôlent les dimensions spectrales de leur parole. Par exemple, des retards dans le retour auditif peuvent entraîner une augmentation de l'intensité (Black, 1951 ; Atkinson, 1953 ; Yates, 1963 ; Howell, 1990) ou une parole « monotone », c'est-à-dire une réduction de la variabilité de la hauteur de la voix (cf. Howell, 2004 ; Chesters et al., 2015).

### **1.2.2 Intégration du retour auditif et somatosensoriel**

Bien que la fonction du retour auditif semble prédominante dans le contrôle moteur de la parole, le retour somatosensoriel est également indispensable pour la précision de la production de la parole. Les informations somatosensorielles doivent également être prédites par le mécanisme prédictif interne pendant la production de la parole, puis comparées aux informations somatosensorielles provenant du retour somatosensoriel. Si une divergence est détectée entre les informations prédites et reçues, le système moteur lance les commandes de corrections nécessaires. Par exemple, une étude réalisée par Jones et Munhall (2003) a démontré qu'une protrusion de la mâchoire induite artificiellement à l'aide d'une prothèse dentaire pouvait avoir un impact significatif sur la production du son /s/. La production de /s/ des locuteurs est revenue à un niveau normal à mesure qu'ils acquéraient de l'expérience avec la prothèse.

Les résultats de l'étude de Jones et Munhall (2003) ont également montré que le retour somatosensoriel et le retour auditif étaient complémentaires. Bien que l'amélioration de la production de /s/ ait été observée même lorsque le retour auditif était masqué, la production avec la présence du retour auditif était encore plus proche de la production normale des locuteurs. En d'autres termes, les informations auditives ont été utilisées pour compenser partiellement la modification du conduit vocal. Par ailleurs, une autre étude menée par Larson, Altman, Liu et Hain (2008) a montré que les locuteurs utilisaient simultanément les informations auditives et somatosensorielles pour contrôler leur parole. Il est donc possible que l'utilisation conjointe de ces deux types de retour explique pourquoi la compensation aux perturbations du retour auditif (ex. le décalage de la  $f_0$  du retour auditif) n'est pas toujours complète lorsque les informations somatosensorielles sont disponibles et normales (Burnett et al., 1998). En effet, lorsqu'un type spécifique du retour est perturbé, tel que le retour auditif, il en résulte une incohérence entre les informations provenant du retour auditif et celles provenant du retour somatosensoriel.

Il est possible d'expliquer la grande variabilité individuelle de l'ampleur de compensation aux perturbations du retour auditif par la préférence sensorielle des individus. En effet, une étude menée par Lametti, Nasir et Ostry (2012) a montré que lorsque les informations venant du retour somatosensoriel et du retour auditif étaient en compétition, certains locuteurs avaient une préférence stable soit pour le retour somatosensoriel soit pour le retour auditif en termes d'ampleur de compensation aux deux types de perturbations (somatosensorielles et auditives) simultanément appliquées pendant la production de la parole. Autrement dit, les locuteurs dépendaient davantage des informations somatosensorielles ou auditives pour stabiliser leur parole.

## **2. Contrôle sensorimoteur de la production de la parole interagissant avec la complexité linguistique**

Au sein du chapitre I, nous avons exposé comment la structure prosodique spécifique à chaque langue contribue à l'émergence de patrons rythmiques qui influencent la régulation du contrôle moteur durant la production de la parole. Dans la première section de ce chapitre, nous avons démontré l'influence des retours sensoriels, tels que le retour auditif, sur le contrôle moteur. Par conséquent, il est pertinent d'explorer l'interaction de ces trois facteurs, à savoir la prosodie/le rythme, le contrôle moteur et le retour auditif, ainsi que d'analyser comment le rythme interagit avec les processus du contrôle sensorimoteur pendant la production de la parole.

### **2.1 Contrôle sensorimoteur modulé par l'expérience linguistique**

Des études ont démontré que l'expérience linguistique des locuteurs influence le contrôle sensorimoteur lors de la production de la parole. Par exemple, la connaissance des locuteurs concernant la catégorie phonémique contribue au contrôle sensorimoteur indépendamment du retour auditif. Dans une étude menée par Niziolek et Guenther (2013), des niveaux de compensation et des réponses corticales ont été comparés suite à un décalage imprévisible des formants du retour auditif appliqué à deux types de voyelles. Les voyelles du premier type étaient situées près de la frontière des catégories phonémiques, tandis que celles du deuxième type étaient situées loin de cette frontière. Le décalage des formants a modifié l'identité phonémique des voyelles du premier type, mais il a affecté seulement les différences perceptives au sein d'une même catégorie phonémique pour les voyelles du deuxième type. Les résultats ont montré que les niveaux de compensation et de réponses corticales des participants étaient plus élevés lorsque le décalage était appliqué près des frontières phonémiques.

De même, Mitsuya et al. (2013) ont mené une étude comparative sur des locuteurs de langues maternelles différentes et ont observé un effet de l'expérience linguistique sur le contrôle du retour auditif. Lorsque les locuteurs francophones et anglophones ont été invités à produire la même voyelle /ε/, la fréquence du deuxième formant du retour auditif a été baissée, ce qui correspond à un arrondissement de l'articulation dans la production du son cible. Les résultats ont montré que

les participants francophones ont compensé les perturbations plus efficacement que les participants anglophones, car la langue française contient des voyelles arrondies qui sont similaires à la voyelle perturbée. En outre, les participants francophones ont également modifié leur troisième formant, ce qui contribue également à l'arrondissement, contrairement aux participants anglophones. Ainsi, la relation entre la voyelle (c'est-à-dire la cible acoustique) et l'espace vocalique spécifique à chaque langue influence le comportement sensorimoteur des locuteurs pour atteindre la cible.

## 2.2 Rôle de l'accentuation dans le contrôle sensorimoteur

Deux études distinctes ont fourni des preuves empiriques soutenant l'hypothèse selon laquelle l'accentuation (cf. chapitre I [2.1](#)) a un impact sur le contrôle sensorimoteur de la parole.

Une étude menée par Kalveram et Jancke (1989) a observé que les locuteurs germanophones allongeaient davantage les syllabes accentuées que les syllabes non accentuées lors de la production d'un mot trisyllabique sous l'effet d'un retour auditif retardé (DAF). De même, l'étude menée par Oschkinat et Hoole (2022) a mis en œuvre des perturbations en temps réel du retour auditif, dans lesquelles l'attaque complexe des consonnes CC était étirée et le noyau vocalique V était comprimé. Lorsque ces perturbations étaient appliquées de manière ciblée aux syllabes non accentuées en début de mot et aux syllabes accentuées en milieu de mot, les germanophones ont accordé une priorité absolue à l'allongement de la durée de la voyelle accentuée afin de compenser les perturbations. Plus précisément, les locuteurs ont allongé localement la voyelle accentuée perturbée, tandis que tous les autres segments dans le mot ont été réduits à une courte portion. De plus, ils ont allongé tous les segments suivant la voyelle non accentuée perturbée pour maintenir constantes les durées relatives dans le mot.

Les deux études susmentionnées suggèrent que les syllabes accentuées des énoncés se caractérisent par une plus grande flexibilité, c'est-à-dire qu'elles sont plus adaptables, dans le cadre du contrôle sensorimoteur. En d'autres termes, dans la condition des perturbations temporelles, la différence de proéminence (en l'occurrence la durée) entre la syllabe accentuée et les autres syllabes au sein du mot devient plus marquée. La syllabe accentuée, en tant que tête de l'unité (le mot isolé dans le cas des deux études susmentionnées), se transforme en un point d'ancrage perceptif plus robuste pour le groupement des syllabes au sein de l'unité. En somme, la prosodie semble jouer un rôle dans le contrôle sensorimoteur lors de la production de la parole, et les patrons rythmiques

(modulés par la prosodie) peuvent ainsi être réorganisés dans la condition des perturbations temporelles pour garantir la stabilité des processus de la production de la parole.

Cependant, il est essentiel de noter que ces deux études n'ont pas réussi à distinguer l'effet de l'accentuation métrique de l'effet de la durée intrinsèque des syllabes accentuées. Par conséquent, les résultats de ces études ne permettent pas de conclure que la plus grande flexibilité des syllabes accentuées en réponse aux perturbations temporelles est uniquement due à leur statut métrique. Il est possible que l'allongement plus prononcé des syllabes accentuées résulte simplement de leur durée physique plus longue. En effet, lorsque des perturbations temporelles sont appliquées en temps réel de manière ciblée, nous ne pouvons pas exclure l'hypothèse qu'un ajustement plus marqué des syllabes accentuées serait essentiellement attribuable au fait que les locuteurs disposent de davantage de temps pour réagir à ces perturbations.

Par ailleurs, ces études se sont concentrées sur des mots prononcés de manière isolée, dans lesquels le mot correspondait à une unité équivalente à un mot prosodique (Astésano, 2017) ou à un syntagme accentuel (Jun et Fougeron, 2000). Il serait pertinent d'approfondir l'analyse du rôle de l'accentuation au sein des énoncés plus étendus, dans lesquels l'accentuation contribue à mettre en place le rythme.

### **2.3 Interaction entre organisation rythmique et contrôle sensorimoteur en production de la parole : perspectives**

Les études examinant le rôle de l'accentuation mettent en lumière la possible contribution organisatrice de la prosodie dans le contrôle sensorimoteur de la production de la parole. Cependant, il est important de noter que très peu d'études se sont directement penchées sur l'exploration de l'interaction entre les patrons rythmiques de surface, tels que le groupement rythmique et la régularité du groupement, et le contrôle sensorimoteur pendant la production de la parole.

Toutefois, comme exposé précédemment dans la section [3.2.2](#) du chapitre I, les patrons rythmiques peuvent s'adapter de manière flexible à la situation actuelle, en fonction de divers facteurs tels que l'interlocuteur ou l'intention de communication. Le groupement rythmique peut également être organisé en fonction de la tâche de production requise ou de la situation de communication. Par exemple, lorsque les locuteurs émettent une séquence de sons de parole à un

débit rapide, des ajustements fréquents sont observés dans la manière dont les événements phonétiques consécutifs sont groupés au sein d'une syllabe (Stetson, 1951 ; Tuller et Kelso, 1991) ainsi qu'entre des syllabes consécutives (Rochet-Capellan et Shwartz, 2007 ; Lancia et Rosenbaum, 2018 ; Lancia et al., 2023). De plus, une étude menée par Abney et al. (2014) a mis en évidence que le groupement des temps de début acoustique (*acoustic onset times*) à différentes échelles temporelles était plus marqué lors de conversations conflictuelles caractérisées par une alternance plus rigide des tours, par rapport aux conversations informelles et amicales.

Dans la section 3 du chapitre I, nous avons avancé la notion de considérer le rythme comme un système coordinatif, où les composantes à différentes échelles temporelles s'harmonisent de manière variable en fonction des caractéristiques propres à la langue et à la situation de communication. Par conséquent, il est particulièrement pertinent d'explorer la manière dont ce système coordinatif s'ajuste en réponse à divers contextes liés au contrôle sensorimoteur de la production de la parole.

### **3. Rôle du sens d'agentivité dans le contrôle sensorimoteur pendant la production de la parole**

Dans la vie quotidienne, nous sommes souvent en train d'interagir avec une ou plusieurs personnes, ce qui signifie que nous sommes constamment exposés à la fois à notre propre voix et à celles des autres. La façon dont le système sensorimoteur d'un locuteur traite la voix perçue pendant la production de la parole diffère selon qu'elle soit produite par le locuteur lui-même (endogène) ou par son interlocuteur (exogène). Par conséquent, notre système sensorimoteur doit constamment s'adapter en fonction des voix perçues.

#### **3.1 Effet des entrées externes sur la production de la parole**

##### **3.1.1 Convergence phonétique**

Il est bien établi que lorsqu'il s'agit d'interactions verbales, les interlocuteurs ont tendance à converger à différents niveaux linguistiques, englobant la structure syntaxique, la forme lexicale et

les caractéristiques phonétiques (Pickering et Garrod, 2004). Dans le domaine de la phonétique, le phénomène de la « convergence phonétique » a fait l'objet d'études dans divers contextes interactifs, notamment les conversations (Pardo, 2006), les tâches de « *shadowing* » où les locuteurs sont invités à répéter immédiatement un stimulus acoustique (ex. l'enregistrement d'un mot) (Pardo et al., 2013), et la parole synchronisée, au cours de laquelle les locuteurs produisent simultanément le même contenu verbal (Cummins, 2002 ; Bradshaw et McGettigan, 2021). Plusieurs paramètres phonétiques entrent en jeu dans la convergence phonétique, par exemple, la fréquence fondamentale (Aubanel et Nguyen, 2020), les formants et la cinématique des articulateurs (Gentilucci et Bernardis, 2007, Mukherjee et al., 2018).

L'interaction entre la prosodie spécifique à chaque langue et le phénomène de convergence phonétique a été observée dans plusieurs études. Par exemple, dans une tâche de « *shadowing* », Michelas et Nguyen (2011) ont démontré que les locuteurs français avaient tendance à produire davantage de tons hauts initiaux (correspondant à un accent initial en français, cf. chapitre I [2.2.2](#)) lorsque le mot stimulus contenait à la fois un ton haut initial et un ton haut final, par rapport à la situation où un mot contenant uniquement le ton haut final a été présenté. Un autre exemple concerne l'imitation des tons en cantonais. Dans une étude menée par Lin, Yao et Luo (2021), les locuteurs cantonais ont été exposés à des stimuli dans lesquels les tons de niveau moyen (ton 3) et les tons de niveau bas (ton 6), qui sont sujets à confusion chez les locuteurs cantonais contemporains, étaient produits de manière distincte, dans une tâche de « *shadowing* ».

### **3.1.2 Cibles sensorimotrices**

Le modèle d'alignement interactif (Pickering et Garrod, 2013) suggère que la convergence est un processus subconscient, activé par un mécanisme d'amorçage automatique. Plus précisément, lorsqu'un individu écoute la parole de son interlocuteur, il imite discrètement la parole de celui-ci, activant ainsi le mécanisme prédictif pour anticiper les énoncés à venir de son interlocuteur sur la base de l'expérience de ses propres actions vocales. Cette imitation cachée tend à activer les mêmes représentations en action dans le système sensorimoteur de l'interlocuteur. Lorsque les prédictions ne correspondent pas aux signaux produits par l'interlocuteur, ces représentations sont modifiées dans le but d'améliorer la prédiction de son comportement. Selon le modèle de Pickering and Garrod (2004), les phénomènes imitatifs sont ainsi dus au changement des représentations des cibles sensorielles d'un individu pendant la perception de la parole de l'interlocuteur. En d'autres termes, lors des interactions verbales, l'entrée externe (c'est-à-dire la voix d'une autre personne)

entraîne un changement de la cible sensorielle (c'est-à-dire les conséquences sensorielles prédites) et une adaptation du système sensorimoteur, ce qui peut donner lieu à une imitation de la production entre les interlocuteurs.

En somme, durant la production de la parole, le système sensorimoteur vise les représentations internes comme cibles. Toutefois, en présence d'entrées externes lors des interactions verbales, les cibles sensorimotrices peuvent être modifiées, conduisant ainsi à l'imitation de ces entrées externes par les locuteurs. Le système sensorimoteur démontre ainsi une flexibilité lui permettant d'ajuster les cibles en fonction d'un environnement en constante évolution.

## **3.2 Génération du sens d'agentivité implicite**

### **3.2.1 Définition du sens d'agentivité**

La capacité d'un individu à exercer un contrôle intentionnel sur ses mouvements articulatoires est étroitement liée au concept d'agentivité. La perception de cette capacité de contrôle est associée à la notion de sens d'agentivité, qui joue également un rôle crucial dans les processus sensorimoteurs impliqués dans la production de la parole.

Le sens d'agentivité est un état mental et neuronal qui nous permet de ressentir (percevoir) le contrôle que nous exerçons sur nos actions et sur le monde extérieur qui en découle (Haggard et Chambon, 2012 ; Haggard, 2017). Cette sensation est liée à plusieurs facteurs, tels que l'intention d'agir, le choix d'une action particulière, l'initiation de l'action, l'exécution de l'action ou encore la réception du retour sensoriel généré par l'action et l'observation des transformations de l'environnement qui en découlent. Le sens d'agentivité est important pour nous car des anomalies ont été constatées chez les patients atteints de troubles mentaux (caractérisés par un sens d'agentivité non conforme aux normes attendues) et il est crucial pour la notion de responsabilité de nos actions (Haggard et Chambon, 2012).

Bien que nous n'y pensions pas souvent dans notre vie quotidienne, nous avons généralement un sens implicite de l'agentivité, qui peut être révélé par des distorsions de la perception du temps (Haggard, 2017). Haggard, Clark, and Kalogeras (2002) ont identifié un phénomène appelé l'effet de « liage intentionnel », qui met en évidence que l'intervalle de temps entre une action volontaire et ses conséquences est perçu comme plus court par l'individu à l'origine de cette action. Cet effet

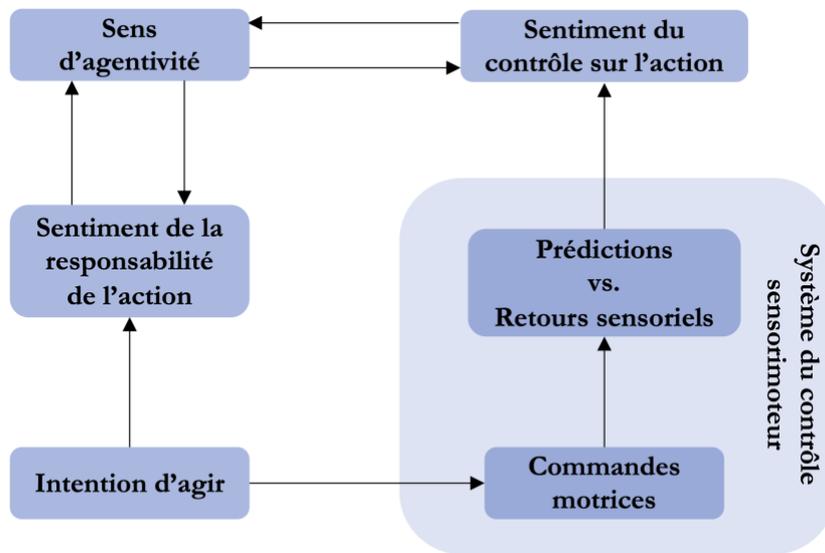
n'est pas observé lorsque l'action est exogène. La compression de la perception de l'intervalle entre l'action et son résultat peut être utilisée comme une mesure permettant de quantifier le sens d'agentivité implicite.

Il convient de souligner que cette forme du sens d'agentivité implicite diffère du jugement explicite d'agentivité, qui peut être mesuré en demandant directement aux individus s'ils ont causé l'action. Le sens d'agentivité explicite est plus vulnérable aux biais cognitifs (Ohata et al., 2022). Ces deux formes d'agentivité ne sont pas toujours cohérentes (Dewey et Knoblich, 2014). Par exemple, l'étude de Zheng et al. (2011) ont démontré que les locuteurs ont manifesté une tendance à identifier une voix inconnue comme étant la leur (modifiée) lorsque cette voix était congruente avec leur attente, dans le contexte où on leur demandait de déterminer si cette voix était la leur. Cependant, simultanément, ils ont imité la voix inconnue en la traitant comme la voix d'une autre personne.

### **3.2.2 Lien entre le sens d'agentivité implicite et le contrôle sensorimoteur**

En considérant l'idée que le sens de l'agentivité (en particulier implicite), est associé à la cohérence entre une action volontaire et ses résultats, il est raisonnable de supposer l'existence d'un lien étroit entre la génération du sens de l'agentivité et les processus du contrôle sensorimoteur. La Figure 7 résume cette relation. L'intention d'agir d'un individu engendre des commandes motrices dans le système de contrôle sensorimoteur. Au sein de ce même système, les retours sensoriels, correspondant aux résultats de l'exécution des commandes motrices, sont comparés aux prédictions, c'est-à-dire les conséquences attendues de l'exécution des commandes motrices (cf. chapitre II 1.1). Le sentiment de contrôle sur l'action émerge lorsque peu ou aucune divergence n'est observée entre les retours sensoriels et les prédictions. Ce sentiment de contrôle sur l'action engendre le sens de l'agentivité. En revanche, si les retours sensoriels perçus ne correspondent pas aux prédictions, le sens de l'agentivité peut être diminué, voire absent (Farrer et al., 2008). D'autre part, l'intention d'agir est la prémisse de l'émergence du sentiment de responsabilité de l'action (c'est-à-dire le sentiment d'être responsable des résultats d'une action), lequel favorise la génération du sens de l'agentivité. Ce dernier renforce le sentiment de l'individu quant au contrôle de l'action endogène. Dans une certaine mesure, la génération du sens de l'agentivité dépend des processus du contrôle sensorimoteur, tout en modulant en retour le contrôle sensorimoteur.

Figure 7 Schéma représentant le lien entre le sens d'agentivité et le contrôle sensorimoteur



À titre d'illustration, dans l'étude menée par Blakemore et al. (2000), les participants ont été sollicités pour évaluer la sensation de chatouillement provoquée par un stimulus tactile appliqué sur la paume de leur main droite. Ce stimulus tactile était généré par des mouvements sinusoïdaux d'une mousse attachée à un bras robotique. Les participants avaient la possibilité de contrôler le bras robotique à l'aide de leur main gauche pour appliquer les stimuli tactiles sur la paume de leur main droite. Différents niveaux de retard et de changement de trajectoire entre les mouvements de la main gauche et l'application des stimuli tactiles sur la paume de la main droite ont été introduits par le bras robotique. Ces interventions du bras robotique engendraient ainsi une divergence entre le retour sensoriel attendu et le retour sensoriel réel. Les résultats de cette étude ont révélé que la sensation de chatouillement augmentait à mesure que cette divergence s'accroissait.

Dans cet exemple, l'intervention du bras robotique dans les mouvements de la main gauche des individus diminue leur sentiment de contrôle sur l'action. La hausse de la sensation de chatouillement correspondrait donc à une diminution du sens d'agentivité. Cela est similaire à notre expérience commune, où nous ressentons une sensation de chatouillement plus intense lorsque quelqu'un d'autre nous chatouille, par opposition à lorsque nous nous chatouillons nous-mêmes. À mesure que l'intervention du bras robotique devient plus intense, les individus perdent progressivement leur sensation de responsabilité concernant les mouvements du bras robotique.

### **3.2.3 Génération du sens d'agentivité implicite pendant la production de la parole**

Pendant la production de la parole, tout comme lors de toute autre activité motrice, la génération du sens d'agentivité est cruciale. La partie A de la Figure 8 illustre ce processus dans différentes conditions de production de la parole. Tout d'abord, lorsque les locuteurs s'expriment dans une condition normale (par exemple, une situation sans perturbations du retour auditif) (cf. la boîte bleue de la partie A de la Figure 8), les retours auditifs perçus doivent correspondre à leurs attentes, ce qui entraîne une suppression de l'activité du cortex auditif (Christoffels et al., 2007 ; Curio et al., 2000 ; Houde et al., 2002 ; Chang et al., 2013). Cette suppression est liée à l'action de parler produite par les locuteurs, car lorsque ces derniers écoutent des stimuli acoustiques qui ont exactement les mêmes caractéristiques, par exemple des enregistrements acoustiques de leur propre production, le phénomène de suppression n'est pas observé. Par conséquent, la suppression de l'activité du cortex auditif joue un rôle essentiel dans la génération du sens d'agentivité. En fait, cette suppression permet aux locuteurs de distinguer entre les sensations endogènes et exogènes (Chang et al., 2013). Cela rejoint l'exemple que nous avons évoqué dans la section 3.2.2 de ce chapitre, où la sensation de chatouillement est réduite lorsque le sujet contrôle lui-même l'action de chatouiller (Blakemore et al., 2000). Le fait que les sons produits par les locuteurs eux-mêmes entraînent la suppression des réponses du cortex auditif aide les locuteurs à mieux se concentrer sur les sons provenant de l'environnement, en particulier dans le contexte de l'interaction verbale. En d'autres termes, la capacité à distinguer entre les sons endogènes et exogènes module l'action de parler des locuteurs.

En revanche, lorsque le retour auditif des locuteurs est perturbé (cf. la boîte orange dans la partie A de la Figure 8), par exemple par un décalage de la  $f_0$  du retour auditif, les retours auditifs perçus par les locuteurs ne correspondent pas à leurs attentes (cf. chapitre II 1.2). Dans ce cas, au lieu de la suppression de l'activité du cortex auditif, l'activité dans ce cortex est renforcée (Behroozmand et al., 2009 ; Chang et al., 2013). Il est important de noter que ce renforcement est particulièrement marqué lorsque les locuteurs parlent en présence de perturbations, par rapport à la condition d'écoute que nous avons mentionnée précédemment. Par conséquent, ce renforcement est également lié à l'action de parler. Le niveau d'activité du cortex auditif suit donc l'ordre suivant :

« parler avec perturbations > écouter > parler normalement »

Sur la base de ce contexte, il est raisonnable de faire l'hypothèse que le renforcement de l'activité du cortex auditif serait associé à une diminution du sens de l'agentivité des locuteurs concernant la

voix produite par eux-mêmes. Cette diminution pourrait conduire à une ambiguïté chez les locuteurs quant à l'origine des sons perçus pendant la production de la parole, ce qui pourrait entraîner des modifications dans leur production (cf. Figure 8 Partie A).

Par ailleurs, nous avançons **l'hypothèse** que le rôle du sens d'agentivité pourrait varier en fonction de l'ampleur des perturbations, comme illustré dans la partie B de la Figure 8. En effet, il a été observé que le renforcement de l'activité du cortex auditif était plus prononcé en réponse à de petites perturbations du retour auditif, telles qu'un faible décalage de la  $f_0$  dans le retour auditif, par rapport aux perturbations plus importantes (Behroozmand et al., 2009). De plus, il a été constaté que plus ce renforcement était fort, plus il s'est révélé prédictive de la compensation motrice (cf. chapitre II section 1) face aux perturbations du retour auditif (Chang et al., 2013). Nous pourrions donc supposer que le renforcement accru de l'activité du cortex auditif serait associé à une diminution moins marquée du sens d'agentivité, car une divergence plus importante entre le retour sensoriel et les prédictions devrait réduire davantage le sens d'agentivité. De plus, nous formulons **l'hypothèse** qu'une diminution plus marquée du sens d'agentivité entraînerait une adaptation moins importante de la production de la parole aux perturbations (cf. également chapitre II 3.4).

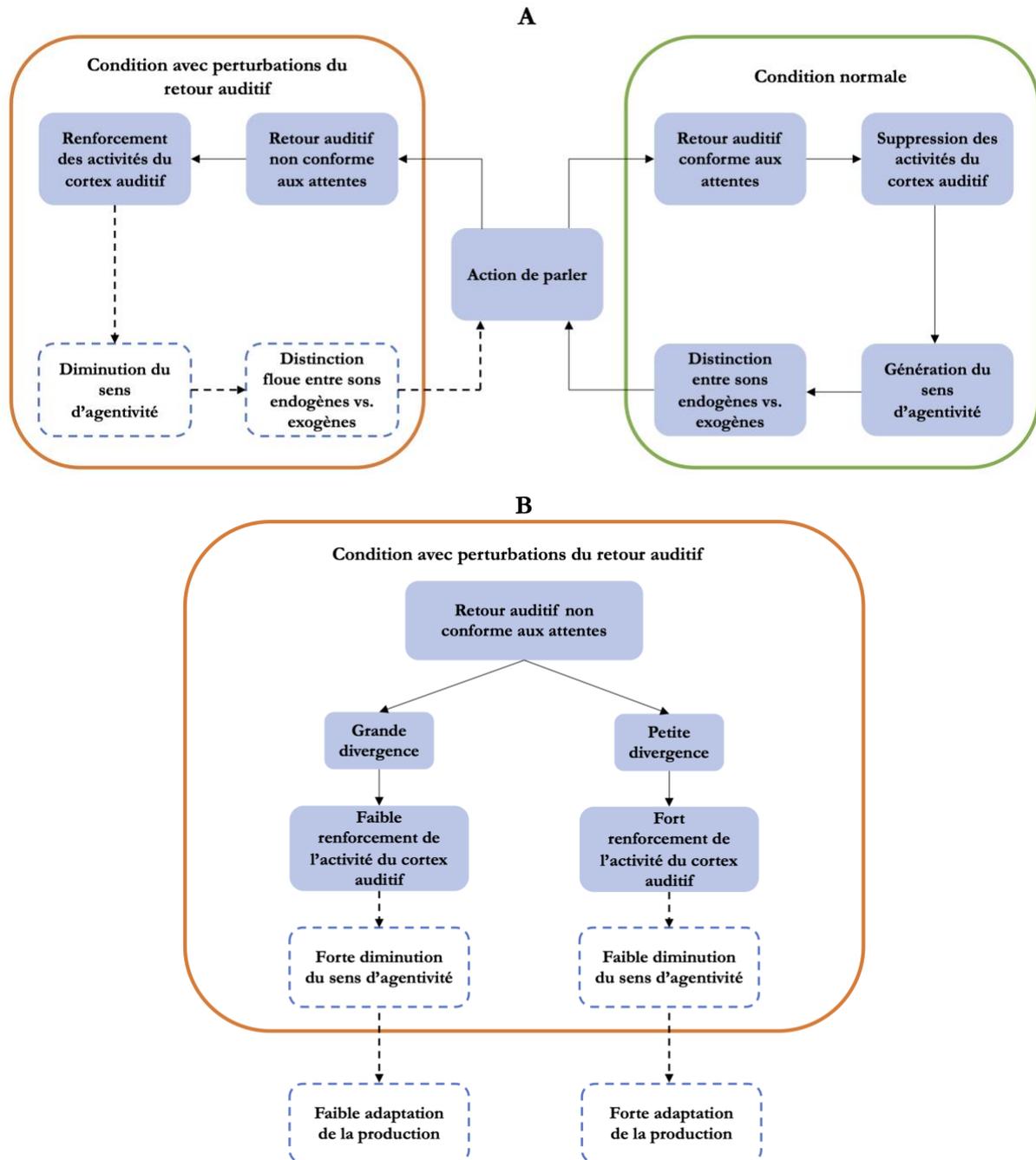
Ces deux hypothèses permettent, par exemple, d'apporter une explication aux résultats obtenus par Liu et Larson (2007) qui démontrent que l'ampleur de l'adaptation aux perturbations ne varie pas en fonction de l'ampleur de ces perturbations. En effet, les perturbations de grande ampleur pourraient induire une réduction plus significative du sens d'agentivité des locuteurs relatif à leur propre voix, ce qui diminuerait leur adaptation aux perturbations de grande ampleur.

Nous tenons à souligner deux points importants concernant ces deux hypothèses. Tout d'abord, il est essentiel de noter que le renforcement de l'activité du cortex auditif et la suppression de l'activité du cortex auditif sont deux mécanismes indépendants, impliquant des activations de différentes zones du cortex auditif (Chang et al., 2013). Par conséquent, un renforcement moins prononcé ne signifie pas nécessairement que l'état de l'activité neuronale dans le cortex auditif se rapproche de celui observé dans la condition de la production normale de la parole. Deuxièmement, il convient de noter que deux types d'adaptation aux perturbations du retour auditif ont été observés dans des études expérimentales. Dans la plupart des cas, les locuteurs optent pour une adaptation de type compensation (cf. chapitre II 1.2.1) afin de réduire la divergence entre les résultats observés de la production et leurs attentes (ex. Burnett et al., 1998). Un second type

d'adaptation consiste à adopter un comportement imitatif (c'est-à-dire suivre les perturbations). Ce mode d'adaptation est moins fréquemment observé dans le cas des perturbations du retour auditif, car l'imitation conduit à une augmentation de la divergence entre les résultats de l'action et les attentes. Cependant, dans certaines situations, notamment lorsque cette divergence est considérable (c'est-à-dire les perturbations sont de grande ampleur), les comportements imitatifs sont plus fréquemment observés (Burnett et al., 1998). Ce sujet sera approfondi dans la section [3.3.1](#) de ce chapitre. Dans le cadre des hypothèses avancées dans cette section, l'expression « faible adaptation » (dans la Figure [8](#)), par exemple, renvoie au type d'adaptation de compensation. En effet, selon nos hypothèses, en cas de forte diminution du sens d'agentivité, les comportements imitatifs devraient être plus fréquemment observés, du fait que la forte réduction du sens d'agentivité pourrait accroître la probabilité pour les locuteurs de percevoir leur voix comme étant d'origine externe.

Figure 8 Illustration du lien entre l'activité du cortex auditif et la génération du sens d'agentivité, ainsi que nos hypothèses concernant le rôle du sens d'agentivité dans la régulation de la production de la parole

Dans la partie A, nous illustrons la génération du sens d'agentivité dans deux conditions de la production de la parole. La partie B se concentre spécifiquement sur deux cas de la condition avec perturbations du retour auditif, mettant en évidence notre hypothèse concernant l'effet de la variation du sens d'agentivité sur la production de la parole en réponse aux perturbations. Les processus marqués par des lignes en pointillés ou un cadre en pointillés correspondent à nos hypothèses.



### 3.3 Flexibilité du sens d'agentivité implicite

#### 3.3.1 Compétition entre les cibles externes et internes

La présence d'une entrée externe peut altérer le sens d'agentivité relatif à notre propre voix. Une étude menée par Jasmin et al. (2016) a démontré que le phénomène de suppression de l'activité du cortex auditif, qui est induit par la production de la parole (cf. chapitre II [3.2.3](#)), n'a pas été observé lors d'une tâche de parole synchronisée, au cours de laquelle deux locuteurs ont été invités à répéter simultanément les mêmes phrases. La parole synchronisée est une forme « extrême » d'interaction verbale qui peut être observée dans notre vie quotidienne. Pendant la parole synchronisée, les locuteurs perçoivent deux voix ayant exactement le même contenu et arrivant à peu près simultanément (avec un décalage temporel subtil). Au cours de la parole synchronisée, le cortex auditif traite la voix des locuteurs eux-mêmes comme s'il s'agissait de la parole d'une autre personne (Jasmin et al., 2016). En d'autres termes, les locuteurs pourraient perdre le sens d'agentivité relatif à leur propre voix.

En dépit de nos connaissances actuelles sur le système sensorimoteur impliqué dans la production de la parole, il reste encore à comprendre comment notre cerveau distingue les cibles sensorimotrices internes et externes. Dans la section [1](#) de ce chapitre, nous avons évoqué la manière dont les informations véhiculées par le retour auditif sont comparées aux prédictions afin de détecter et corriger les erreurs. Afin de rectifier les erreurs, le système moteur émet des commandes de compensation. Par exemple, lorsque la  $f_0$  produite est plus élevée que prévue, la  $f_0$  produite lors de la prochaine production sera abaissée.

Cependant, certains individus ne répondent pas toujours de manière compensatoire aux perturbations de leur retour auditif, comme l'ont montré plusieurs études (Burnett et al., 1998 ; Hain et al., 2000 ; Larson et al., 2007 ; Behroozmand et al., 2012 ; Franken et al., 2018). Ces études ont révélé que certains locuteurs modifiaient leur production en suivant la direction du décalage de la  $f_0$  de leur retour auditif, même si cela conduisait à une divergence entre la voix produite et la voix prédite. Bien qu'il puisse être difficile pour les locuteurs de distinguer la direction du décalage de la  $f_0$  (Neuhoff et al., 2002), notamment lorsque le décalage est imprévisible et de petite amplitude, des études ont révélé que la réponse suivant la direction du décalage était plus importante lorsque le décalage était prévisible (Behroozmand et al., 2012) ou plus important (Burnett et al., 1998). Par conséquent, la difficulté à distinguer la direction du décalage de la  $f_0$  ne pourrait pas être le seul

facteur contribuant aux réponses suivant le décalage. Il est également possible que le sens d'agentivité des locuteurs relatif à leur propre voix soit altéré par le décalage de la  $f_0$  dans leur retour auditif, les amenant temporairement à traiter la voix perçue comme étant générée de l'extérieur (cf. chapitre II [3.1.2](#)). Toutefois, les réponses compensatoires restent prédominantes même chez les locuteurs qui suivent la perturbation dans certains essais expérimentaux (Burnett et al., 1998), suggérant une « compétition » entre les cibles sensorimotrices internes et externes pendant le contrôle sensorimoteur ainsi qu'une perturbation du sens d'agentivité.

### **3.3.2 Altération du sens d'agentivité implicite**

Il a été constaté que perturber la  $f_0$  du retour auditif peut entraîner une diminution du sens d'agentivité implicite chez les locuteurs. Par exemple, dans une récente étude menée par Ohata et al. (2022), les participants ont été invités à estimer l'intervalle entre leur production articulatoire et leur retour auditif artificiellement retardé. Les résultats ont démontré que lorsque la  $f_0$  du retour auditif était décalée vers le haut ou vers le bas, les participants avaient tendance à estimer cet intervalle comme étant plus long par rapport à la condition où la  $f_0$  du retour auditif n'était pas altérée. En d'autres termes, le décalage de la  $f_0$  du retour auditif des participants a réduit l'effet de « liage intentionnel » décrit dans la section [3.2.1](#) de ce chapitre, suggérant ainsi une diminution du sens d'agentivité implicite chez les participants.

La perturbation continue de la  $f_0$  du retour auditif des locuteurs pourrait également altérer leur sens d'agentivité, car elle est souvent associée à des modifications de la cible du contrôle sensorimoteur. Par exemple, l'étude de Franken et al. (2022), a observé que les locuteurs alignaient leur  $f_0$  sur la  $f_0$  du retour auditif continuellement décalé (+500 cents ou -500 cents), ce qui suggère que le retour auditif continuellement perturbé pourrait être perçu comme étant généré à l'extérieur. De plus, une exposition prolongée à un décalage continu de la  $f_0$  du retour auditif a entraîné un changement temporaire de l'attente des locuteurs de leur  $f_0$ , dans la direction du décalage constant de la  $f_0$ . Par exemple, après avoir été exposés à un décalage continu et positif de la  $f_0$  (d'une grande ampleur de +500 cents) du retour auditif, les locuteurs étaient moins susceptibles de juger que la  $f_0$  de leur retour auditif était différente de celle attendue, même lorsque la  $f_0$  de leur retour auditif était positivement décalée de manière imprévisible de petite ampleur (30, 60, 100 ou 150 cents). Cet effet résiduel confirme le basculement des représentations internes résultant de l'exposition prolongée à un décalage continu de la  $f_0$  du retour auditif. Un effet résiduel similaire « à court terme » (Sato et al., 2013) ou « à long terme » (Pardo et al., 2012) a été observé dans des études sur la

convergence phonétique, dans lesquelles les locuteurs alignent leurs caractéristiques acoustiques sur celles des entrées vocales externes (venant d'un stimulus acoustique ou d'une autre personne).

Le basculement des cibles sensorimotrices ne peut pas simplement être expliqué par la continuité de la perturbation. Dans une étude menée par Franken et al. (2019), la perturbation continue des formants du retour auditif a entraîné une compensation plus forte de la perturbation plutôt qu'un basculement des cibles sensorimotrices. Par conséquent, l'alignement de la  $f_0$  sur la  $f_0$  du retour auditif décalé pourrait être le résultat d'un processus dans lequel la  $f_0$  du retour auditif est traitée comme une cible externe.

### **3.4 Sens d'agentivité implicite modulant le contrôle sensorimoteur de la parole**

Il reste des incertitudes quant à la manière dont la réduction du sens d'agentivité et le basculement des cibles sensorimotrices peuvent influencer le contrôle sensorimoteur lors de la production de la parole. Récemment, une étude menée par Franken et ses collègues (2021) a appliqué simultanément un décalage continu de grande ampleur de la  $f_0$  du retour auditif (+500 cents) et des décalages plus courts et imprévisibles de plus petite amplitude (+100, -100, +400, -400 cents). Les participants ont aligné leur  $f_0$  sur la  $f_0$  de leur retour auditif largement décalé, mais ont compensé les petits décalages imprévisibles de la  $f_0$  du retour auditif dans leur production. Les auteurs ont affirmé l'absence de l'altération du sens d'agentivité causée par le décalage continu de la  $f_0$  de grande amplitude, en raison des réponses de compensation observées chez les participants face aux divers niveaux de petits décalages imprévisibles.

Cependant, le fait que les participants ont imité la  $f_0$  de leur retour auditif largement décalé pourrait déjà être un indicateur de la réduction du sens d'agentivité, car ils ont traité la voix du retour auditif comme une cible externe. Nous émettons l'hypothèse que l'exposition à un large décalage de la  $f_0$  du retour auditif pourrait fait basculer la cible sensorimotrice. Lorsque les petits décalages de la  $f_0$  du retour auditif ont été introduits de manière imprévisible, une divergence entre la nouvelle cible sensorimotrice et le retour auditif actuel aurait dû être détectée. Cela pourrait expliquer pourquoi les locuteurs ont continué à compenser les petits décalages de la  $f_0$  de leur retour auditif, alors qu'ils ont suivi le large décalage de la  $f_0$  de leur retour auditif. Dans une certaine mesure, cet exemple confirme la flexibilité de notre cible sensorimotrice.

Il est important de noter que la production de la parole implique différentes dimensions, qui peuvent être traitées par des processus indépendants ou interactifs. Il est donc pertinent de se demander si le basculement de la cible sensorimotrice sur la  $f_0$  peut influencer le contrôle d'autres dimensions de la parole, telles que la dimension temporelle (cf. chapitre II [1.2.1](#)).

## 4. Résumé du chapitre

Malgré sa grande complexité, notre système moteur fait preuve de la flexibilité nécessaire pour s'adapter à diverses conditions au cours de la production de la parole. Cette adaptabilité est rendue possible grâce à l'intégration sensorimotrice, qui englobe la comparaison des retours sensoriels, tels que le retour auditif et le retour somatosensoriel, associés à une action endogène, avec les résultats attendus. Cette comparaison permet un ajustement en temps réel des commandes motrices en fonction de l'état actuel du système sensorimoteur et du contexte de la production de la parole (Parrell et al., 2019). Les prédictions des résultats, générées en fonction des objectifs linguistiques et de l'état courant du système sensorimoteur, peuvent également contribuer à l'ajustement du système sensorimoteur (Wolpert et al., 1995 ; Miall et Wolpert, 1996), en particulier lorsque les retours sensoriels sont retardés ou perturbés.

Les recherches sur la manipulation du retour sensoriel (Burnett et al., 1998 ; Houde et Jordan, 1998, 2002 ; Larson et al., 2007 ; Jones et Munhall, 2003) illustrent la manière dont les retours auditif et somatosensoriel fournissent des informations complémentaires, intégrées dans les processus de contrôle sensorimoteur. Cette thèse se concentre spécifiquement sur le retour auditif, qui englobe plusieurs dimensions, telles que la  $f_0$ , l'intensité et les formants (Burnett et al., 1998 ; Houde et Jordan, 1998, 2002 ; Larson et al., 2007). Lorsque le retour auditif est altéré (ex. lorsque la  $f_0$  du retour auditif est positivement décalée), dans la plupart des cas, une adaptation motrice de compensation se produit (ex. une diminution de la  $f_0$ ). En cas du retard du retour auditif, les locuteurs ralentissent leur débit de parole (Black, 1951) afin de minimiser le décalage perçu entre les mouvements articulatoires et leurs conséquences acoustiques.

Le contrôle sensorimoteur peut être influencé par les caractéristiques linguistiques, telles que la prosodie, spécifiques à chaque langue. Par exemple, les voyelles accentuées se montrent plus flexibles et sensibles à l'adaptation (ex. avec un allongement plus prononcé) face à des perturbations

temporelles du retour auditif (ex. Kalveram et Jancke, 1989). Nous avons formulé l'hypothèse que le rythme de la parole, qui est modulé par la prosodie et qui joue un rôle significatif dans la simplification du contrôle moteur (cf. chapitre I), s'adapterait de manière flexible à diverses conditions de contrôle sensorimoteur.

De plus, lorsqu'ils participent à une interaction verbale, les locuteurs doivent être en mesure de distinguer si la voix qu'ils perçoivent est la leur (endogène) ou celle d'une autre personne, car le système sensorimoteur traite différemment ces deux types de voix. Lorsque la voix perçue provient d'un interlocuteur, le système sensorimoteur ajuste sa cible en fonction de cette voix perçue (Pickering and Garrod, 2004), ce qui entraîne une imitation de la voix de l'interlocuteur (Pardo, 2006 ; Pardo et al., 2013 ; Cummins, 2002 ; Bradshaw et McGettigan, 2021). La capacité à distinguer la source de la voix perçue est associée au sens d'agentivité, qui renvoie à la perception du contrôle d'une action endogène (Haggard, 2017).

La génération du sens d'agentivité est liée à la comparaison entre les retours sensoriels et les attentes dans le contrôle sensorimoteur (Farrer et al., 2008 ; Blakemore et al., 2000). Nous postulons que lorsque cette comparaison génère une divergence, par exemple lorsque la  $f_0$  du retour auditif ne correspond pas à la prédiction concernant la  $f_0$ , cela pourrait réduire le sens d'agentivité des locuteurs relatif à leur propre voix. En fait, lorsque la  $f_0$  du retour auditif est continuellement décalée de manière significative, les locuteurs ont tendance à suivre cette  $f_0$  décalée, comme s'ils percevaient la voix entendue comme celle d'une autre personne (Franken et al., 2021). Nous formulons l'hypothèse que ce changement de la cible sensorimotrice peut être associé à une réduction du sens d'agentivité, et que cette réduction peut moduler le contrôle sensorimoteur lors de la production de la parole.

Nous avons abordé la production de la parole en tant que tâche à la fois complexe (chapitre I) et flexible (chapitre II). Dans le chapitre à venir, nous présenterons nos questions de recherche ainsi que les hypothèses qui y sont associées, en intégrant ces deux aspects de la production de la parole. De plus, nous détaillerons les méthodes que nous avons mises en place pour explorer ces hypothèses.

---

# CHAPITRE III : MESURER LES PATRONS RYTHMIQUES DE LA PAROLE PRODUITE SOUS PERTURBATION(S) AUDITIVES(S)

## Sommaire du chapitre

<b>1. Questions de recherche et hypothèses.....</b>	<b>54</b>
1.1 Question de recherche I et hypothèses .....	54
1.2 Question de recherche II et hypothèses.....	57
<b>2. Expérimentation.....</b>	<b>59</b>
2.1 Expérience I Retour auditif retardé à différent niveaux (expérience pilote).....	59
2.1.1 Objectives.....	59
2.1.2 Participants.....	60
2.1.3 Matériaux.....	60
2.1.4 Procédure.....	61
2.1.5 Traitement des données.....	62
2.2 Expérience II Partie A – condition uniquement DAF .....	64
2.2.1 Objectives.....	64
2.2.2 Participants.....	64
2.2.3 Matériaux.....	64
2.2.4 Procédure.....	65
2.2.5 Traitement des données.....	66
2.3 Expérience II Partie B – condition perturbations multiples .....	66
2.3.1 Objectives.....	66
2.3.2 Participants, matériaux, procédure et traitement des données .....	66
<b>3. Méthodes d’analyses .....</b>	<b>67</b>
3.1 Mesurer la différence d’allongement entre les syllabes accentuées et les syllabes non accentuées sous l’effet du DAF.....	68
3.2 Mesurer la coordination entre la composante syllabique et la composante supra-syllabique ..	71
3.2.1 Présentation générale de la méthode.....	71
3.2.2 Extraction de la composante syllabique et la composante supra-syllabique .....	72
3.2.3 Normalisation d’amplitude .....	73
3.2.4 Calcul de l’indice de coordination – Phase Locking Value (PLV).....	75
<b>4. Résumé du chapitre .....</b>	<b>76</b>

La tâche de la production de la parole se caractérise par une complexité intrinsèque et une flexibilité marquée. Nous avons décrit comment le système de production de la parole gère cette complexité et assure sa flexibilité. D'un côté, le rythme, en tant que système de coordinatif, émerge de la coordination des articulateurs lors de la production de la parole, et contribue à simplifier le contrôle moteur. La génération du rythme est modulée par les structures prosodiques spécifiques

à chaque langue (cf. chapitre I). De l'autre côté, un système intégrant le contrôle moteur, les retours sensoriels et le mécanisme prédictif garantit la flexibilité de la production de la parole. Les processus au sein de ce système peuvent être modulés par des facteurs linguistiques tels que la prosodie (cf. chapitre II section 2), ainsi que par la perception des locuteurs concernant le contrôle de leurs actions d'articulation, c'est-à-dire le sens d'agentivité (cf. chapitre II section 3). Nous proposons donc d'étudier la complexité et la flexibilité de la production de la parole dans une perspective intégrative, en prenant en considération l'interaction entre la prosodie, le rythme qui en découle, et le sens d'agentivité en production de la parole.

Dans ce chapitre, nous commencerons par présenter nos questions de recherche et les hypothèses qui y sont associées (section 1). Nous détaillerons ensuite les expériences des perturbations du retour auditif menées pour explorer ces questions de recherche (section 2). Enfin, nous décrirons nos méthodes d'analyse des données collectées des expériences (section 3).

## **1. Questions de recherche et hypothèses**

### **1.1 Question de recherche I et hypothèses**

Le système de contrôle sensorimoteur garantit la flexibilité de la production de la parole (cf. chapitre II section 1). La prosodie favorise l'émergence de groupements rythmiques (ex. le groupement des syllabes) à diverses échelles temporelles (cf. chapitre I section 2), ce qui réduit la complexité du contrôle moteur et stabilise la production de la parole. Un degré élevé de groupement syllabique est associé à une coordination stable entre la composante syllabique et la composante supra-syllabique dans le rythme (cf. Figure 5). Jusqu'à présent, il a été accordé très peu d'attention à l'étude de la manière dont le groupement rythmique et la coordination entre les composantes rythmiques (cf. chapitre I 3.2.2) interagissent avec les processus du contrôle sensorimoteur. Dans ce contexte, nous formulons notre **première question de recherche** : **Quelle est l'interaction entre la prosodie et les processus du contrôle sensorimoteur, et dans quelle mesure cette interaction influe-t-elle sur l'organisation rythmique au cours des processus du contrôle sensorimoteur ?**

*Hypothèse 1) : La flexibilité du contrôle sensorimoteur est influencée par l'accentuation à différents niveaux prosodiques (cf. chapitre IV).*

Cette première hypothèse repose sur les études existantes ayant examiné les différences de réponses au retard du retour auditif (DAF) lors de la production de syllabes accentuées par rapport à celles des syllabes non accentuées (cf. chapitre II 2.2). Ces recherches ont constaté que les syllabes accentuées étaient plus allongées sous l'influence du DAF. En d'autres termes, l'accentuation rend certaines parties des énoncés plus flexibles pour s'adapter aux perturbations dans les processus de contrôle sensorimoteur. Cependant, comme précisé dans la section 2.2 du chapitre II, les syllabes accentuées sont généralement physiquement plus longues que les syllabes non accentuées, ce qui offrirait aux locuteurs davantage de temps pour réagir au DAF pendant la production des syllabes accentuées par rapport à celle des syllabes non accentuées. Notre hypothèse suggère que la plus grande flexibilité des syllabes accentuées serait associée à leur position dans la hiérarchie prosodique, indépendamment de la durée physique de ces syllabes.

De plus, il est important de noter que les études citées dans la section 2.2 du chapitre II ont principalement restreint leurs investigations à l'accentuation au sein de mots isolés. Cependant, étant donné que l'accentuation est organisée de manière hiérarchique et s'exprime à différents niveaux de proéminence (cf. chapitre I 2.2 et 2.3), la force métrique des niveaux d'accentuation peut varier. Par conséquent, nous formulons l'hypothèse que la flexibilité des syllabes accentuées, associées à différents niveaux prosodiques, peut différer en ce qui concerne le contrôle sensorimoteur. Par exemple, les syllabes accentuées qui délimitent des syntagmes intonatifs pourraient présenter une plus grande flexibilité que celles délimitant des syntagmes accentuels.

Afin d'explorer nos hypothèses (y compris l'hypothèse 1 et les autres hypothèses qui seront exposées dans les paragraphes suivants), nous avons mené des expériences en utilisant le paradigme du retard du retour auditif (DAF). Cependant, nous avons adapté les stimuli et les paramètres des expériences (cf. chapitre III section 2) pour mieux répondre à nos questions de recherche.

Dans le cadre de cette thèse, nous avons choisi de travailler avec des locuteurs francophones comme sujets d'étude (cf. chapitre III section 2). Le système métrique du français se caractérise par sa bipolarité (cf. chapitre I 2.2.2). En d'autres termes, en plus de l'accent final (à la frontière droite des unités prosodiques) qui est métriquement obligatoire et qui sert à délimiter les unités, un accent initial (à la frontière gauche des unités) est également autorisé. Dans cette thèse, notre attention se porte sur le rôle de l'accent final, qui joue un rôle métrique essentiel, dans le contrôle sensorimoteur.

Il convient de noter que le rythme est multidimensionnel. En d'autres termes, plusieurs paramètres, tels que la durée, la fréquence fondamentale et l'intensité, peuvent interagir dans la perception et la production du rythme. Cependant, étant donné que le DAF constitue une perturbation temporelle dans la production de la parole, la variation temporelle des unités linguistiques est directement affectée par le DAF. De plus, la durée représente un paramètre essentiel pour l'accentuation en français (cf. chapitre I [2.2.2](#)). Par conséquent, les effets secondaires du DAF, tels que les variations d'intensité et de  $f_0$  (cf. chapitre II [1.2.1](#)), ne sont pas pris en considération dans le cadre de cette première hypothèse. Néanmoins, ils ont été examinés en vue d'évaluer le sens d'agentivité, en lien avec notre troisième hypothèse (cf. hypothèse [3](#)).

*Hypothèse 2) : Les perturbations du retour auditif, telles que le DAF, induisent une réorganisation de la coordination des composantes rythmiques (cf. chapitre V).*

Conformément à la première hypothèse, les accents jouent un rôle d'ancrage, non seulement dans la perception du groupement des éléments, comme suggéré par Dauer (1983, 1987) (cf. chapitre I [2.5](#)), mais également dans les processus de contrôle sensorimoteur. Face à des perturbations telles que le DAF, nous nous attendons à ce que les différences de proéminence (ex. la durée) entre les syllabes accentuées et non accentuées augmentent, ce qui entraînerait un groupement plus marqué des syllabes, tant sur le plan perceptif, selon l'hypothèse de Dauer (1983, 1987), que sur le plan de la production (notre hypothèse). Il a été démontré qu'un degré important de groupement entre les syllabes était associé à une relation temporelle plus stable entre la composante syllabique et la composante supra-syllabique (cf. chapitre I [3.2.2](#)). Par conséquent, nous formulons l'hypothèse que sous l'effet du DAF, le système de production de la parole requiert une coordination (c'est-à-dire une relation temporelle) plus stable entre la composante syllabique et la composante supra-syllabique afin d'assurer sa stabilité.

Au sein de cette thèse, la relation temporelle entre les composantes rythmiques est mesurée par le degré de synchronisation des phases entre les oscillations d'amplitude syllabiques et les oscillations supra-syllabiques (cf. chapitre III [3.2](#)).

## 1.2 Question de recherche II et hypothèses

Lorsque nous sommes engagés dans une interaction verbale, notre système de contrôle sensorimoteur doit prendre en considération non seulement les retours sensoriels liés à notre propre parole, mais aussi les informations qui proviennent de sources externes, telles que les sons produits par notre interlocuteur. De plus, il convient de noter que la manière dont le système sensorimoteur traite la voix endogène (notre propre voix) diffère de la façon dont il traite la voix exogène (la voix d'autrui) (cf. chapitre II [3.1](#)). Dans cette optique, la capacité d'un locuteur à distinguer la source d'une voix perçue pendant la production de la parole revêt une grande importance. Le sens d'agentivité, qui reflète la perception d'un individu sur son contrôle d'une action endogène (cf. chapitre II [3.2.1](#)), joue un rôle crucial dans cette capacité, ainsi que dans la production de la parole. Le sens d'agentivité émerge des processus du contrôle sensorimoteur, notamment du processus de comparaison entre les retours sensoriels et les attentes, et en retour il module le contrôle sensorimoteur (cf. chapitre II [3.2.2](#)).

En prenant en compte simultanément les deux facteurs - la prosodie et le sens d'agentivité - qui peuvent influencer le contrôle sensorimoteur dans la production de la parole, nous formulons notre **deuxième question de recherche : Comment interagissent le rôle organisateur de la prosodie et celui du sens d'agentivité dans les processus du contrôle sensorimoteur ?**

*Hypothèse 3 : Une perturbation continue de la f0 du retour auditif diminue le sens d'agentivité des locuteurs, et cette diminution se traduit par une réduction des réponses des locuteurs au DAF (cf. chapitre VI).*

Dans la production de la parole, les perturbations du retour auditif créent une divergence entre les retours auditifs et les attentes des locuteurs concernant leur propre voix. Nous supposons que cette divergence réduit le sens d'agentivité des locuteurs (cf. chapitre II [3.2.2](#)).

En effet, l'étude menée par Franken et al. (2022), que nous avons mentionnée dans la section [3.3.2](#) du chapitre II, a révélé que les locuteurs avaient tendance à suivre une perturbation continue de f0 de leur retour auditif. En d'autres termes, ils ont ajusté leur f0 dans la même direction que le décalage artificiellement introduit dans leur retour auditif. Cette adaptation des locuteurs en réponse à la perturbation du retour auditif suggère qu'ils ont traité la voix entendue comme si elle provenait d'une autre personne, de manière similaire au phénomène de la « convergence phonétique » observé dans l'interaction verbale (cf. chapitre II [3.1.1](#)). Autrement dit, la cible du

contrôle sensorimoteur, en l'occurrence la  $f_0$ , serait été modifiée par la perturbation continue de la  $f_0$  du retour auditif.

Nous supposons que si ce phénomène correspond à une réduction du sens d'agentivité, les locuteurs seraient moins perturbés par le DAF, ce qui signifie qu'ils produiraient moins d'adaptations, telles que l'allongement des syllabes, en réponse au DAF. En d'autres termes, lorsque le sens d'agentivité des locuteurs est réduit, ils ont l'impression implicite d'être moins responsables du retard de leur voix (cf. chapitre II [3.2](#)), ce qui diminue leur sentiment de nécessité d'ajuster leur production en réponse au DAF. Dans cette perspective, la réduction du sens d'agentivité régule le contrôle sensorimoteur.

Ainsi, pour explorer notre deuxième question de recherche, nous avons mis en place un paradigme expérimental où les participants ont été soumis à la fois au DAF et à une perturbation de la  $f_0$  de leur retour auditif (cf. chapitre III [2.3](#)).

*Hypothèse 4) : Dans la régulation du contrôle sensorimoteur, le rôle de la prosodie et celui du sens d'agentivité interagissent, et cette interaction modifie l'organisation rythmique (cf. chapitre [VI](#)).*

En nous basant sur l'hypothèse [3](#), nous postulons que la prosodie interviendrait dans la régulation du sens d'agentivité sur le contrôle sensorimoteur. Plus précisément, si la réduction du sens d'agentivité entraîne une diminution de l'ampleur des réponses des locuteurs au DAF, alors la différence de proéminence (la durée) entre les syllabes accentuées et non accentuées serait réduite. Cette différence est supposée augmenter sous l'effet du DAF dans une condition normale (cf. hypothèse [1](#)).

De plus, la réduction du sens d'agentivité devrait également entraîner une diminution de la réorganisation rythmique induite par le DAF. En d'autres termes, lorsque le sens d'agentivité des locuteurs est réduit, ils ressentiraient implicitement moins la nécessité de renforcer la coordination entre les composantes rythmiques pour maintenir la stabilité dans le contexte du DAF.

## 2. Expérimentation

Pour investiguer les questions de recherche que nous avons proposées dans la section 1, nous avons conçu deux expériences de perturbations du retour auditif. La première expérience, bien que préliminaire, joue un rôle essentiel pour nous aider à déterminer les paramètres de la deuxième expérience et à examiner les méthodes d'analyse spécifiquement développées dans le cadre de cette thèse. La deuxième expérience est divisée en deux parties, correspondant à deux conditions de perturbation du retour auditif. La Partie A correspond à la « **condition uniquement DAF** ». Dans cette condition, les sujets n'ont été perturbés que par le DAF. En revanche, la Partie B correspond à la « **condition perturbations multiples** », dans laquelle les sujets ont été perturbés simultanément par le DAF et un décalage continu de la  $f_0$  du retour auditif. La Partie A vise principalement à répondre à notre première question de recherche, tandis que la Partie B vise à répondre à la deuxième question de recherche.

### 2.1 Expérience I Retour auditif retardé à différent niveaux (expérience pilote)

#### 2.1.1 Objectives

Cette première expérience a été conçue avec plusieurs objectifs distincts. En premier lieu, en employant plusieurs niveaux du DAF, notre intention est d'évaluer et comparer les effets induits par ces différents niveaux du DAF. Il a été observé dans des recherches antérieures que l'ampleur des effets du DAF atteint un point de saturation à un certain niveau (Fairbanks, 1955 ; Black, 1951 ; Zimmermann et al., 1988). Cependant, ce niveau de saturation peut varier en fonction de plusieurs facteurs, tels que le débit de parole (Robinson, 1972 ; Stuart et Kalinowski, 2015). Par conséquent, cette expérience vise à déterminer les niveaux de DAF les plus adaptés au type de stimuli utilisés et à nos sujets francophones.

De plus, cette expérience combine des phrases normales avec des virelangues parmi les stimuli. L'inclusion de virelangues vise à évaluer la spécificité des effets du DAF sur la production de la parole. Le DAF introduit en effet un degré de difficulté dans la production de la parole, se manifestant par une augmentation des hésitations et des erreurs d'élocution, similaires à celles observées lors de la production de virelangues. Si le DAF induit une réorganisation rythmique (ex.

un renforcement de la coordination entre les composantes rythmiques) en vue de stabiliser la production, il est pertinent de s'interroger sur l'origine de cette réorganisation, à savoir si elle est imputable à la perturbation temporelle induite par le DAF ou aux disfluences provoquées par le DAF.

Enfin, cette expérience nous permet également d'évaluer les méthodes d'analyse spécifiquement élaborées pour cette étude (cf. chapitre III section 3). Une comparaison des résultats obtenus dans le cadre de cette première expérience avec ceux de la deuxième expérience permettra de valider ces méthodes d'analyse et de confirmer leur pertinence pour étudier les patrons rythmiques de la parole sous l'influence de perturbations auditives.

### **2.1.2 Participants**

L'expérience I a impliqué la participation de six locutrices de langue maternelle française âgées de 22 à 31 ans, avec un âge moyen de 24,8 ans. Toutes les participantes étaient des étudiantes diplômées ou en doctorat de l'Université Sorbonne Nouvelle à Paris et ont accepté de participer volontairement à l'expérience sans rémunération. Aucune des participantes n'a signalé de troubles de la parole ou de la voix, ni de déficience auditive. Toutes les participantes ont signé un accord nous autorisant à utiliser les données collectées pour des analyses et des publications anonymes. Nous avons limité la participation aux locutrices en raison de la sensibilité différente aux perturbations du retour auditif entre les hommes et les femmes, comme cela a été démontré dans des travaux antérieurs (ex. Bachrach, 1964 ; Howell et Archer, 1984).

### **2.1.3 Matériaux**

Les stimuli comprenaient trois phrases en français, dont deux virelangues, générant des patrons fortement rythmiques.

- 1) /valmɔ̃ vuʁ vvar lə vɔl/ (Valmont voulait voir le vol.)
- 2) /lə nɔrmɑ̃ ɔr nɔrm mɔr/ (Le Norman hors norme mord.)
- 3) /mamɑ̃ e mami mani ma mɛ̃/ (Maman et Mamie manient ma main.)

Les trois phrases utilisées dans cette étude ont été conçues pour générer une parole rythmée grâce à la présence des séquences de phonèmes répétées. Cependant, de légères déviations par rapport à ces patrons répétés étaient également introduites pour augmenter la difficulté de la tâche de production de la parole. La première phrase avait un patron d'attaque régulier, commençant

principalement par une structure CV dont l'attaque était /v/. Cette phrase était considérée comme étant la plus facile à prononcer. La deuxième phrase comportait la répétition de la séquence /norm/ suivie de /mor/ (les sons répétés apparaissant dans un ordre inversé). La troisième phrase présentait la répétition de la séquence /mVmV/, mais avec une interruption introduite par la présence de /mVnV/. Les énoncés de haute rythmicité ont été délibérément sélectionnés dans le but de faciliter nos investigations relatives aux patrons rythmiques.

#### **2.1.4 Procédure**

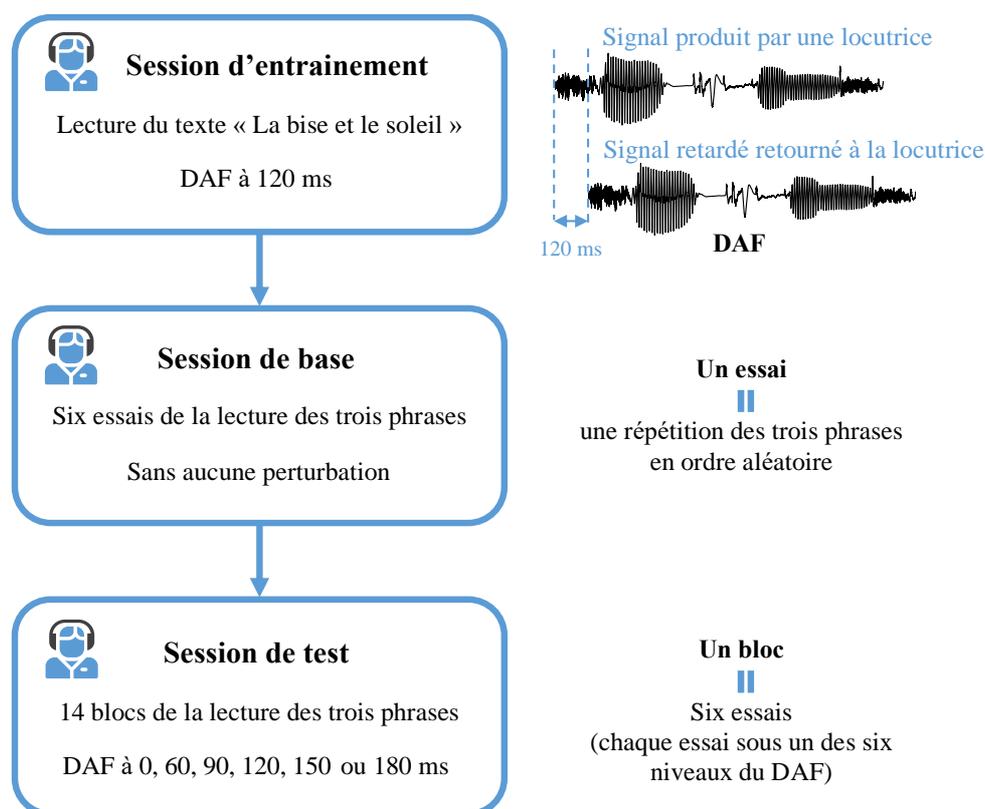
L'expérience a été menée dans une salle anéchoïque spécialement conçue pour éliminer les réflexions sonores et réduire les bruits de fond. Les participantes étaient confortablement installées devant un ordinateur portable sur lequel étaient affichées les instructions relatives à la tâche à accomplir. Les participantes étaient équipées d'un microphone et d'un casque audio pour enregistrer leur voix et entendre leur propre production, avec ou sans retard audio manipulé. Tous les essais ont comporté un petit retard imperceptible dû au traitement audio en temps réel. L'ensemble de l'expérience a été réalisé avec le logiciel MATLAB (The MathWorks Inc., Natick, MA) et la manipulation du retard audio a été effectuée par le biais du programme AUDAPTER (Cai et al., 2008).

La Figure 9 illustre la procédure de l'expérience. Celle-ci a débuté par une session d'entraînement, durant laquelle les participantes ont été invitées à lire un court texte en français intitulé « La bise et le soleil », avec un DAF à 120 ms. Le volume des signaux transmis par le casque a été réglé à un niveau confortable mais suffisamment fort pour minimiser la perception par la locutrice de sa voix non manipulée (Burnett et al., 1998 ; Liu et al., 2012). Au cours de cette session, aucune donnée de production des participantes n'a été enregistrée.

Au cours des deux sessions suivantes, les participantes ont été invitées à lire les trois phrases stimuli, sans interruption et à un rythme qui leur était confortable, dans l'ordre affiché sur l'écran de l'ordinateur portable. Chaque essai, comprenant une répétition des trois phrases, avait une durée prédéfinie de 11 secondes (avec possibilité d'ajustement après les six premiers essais avec DAF en cas de débit de parole particulièrement lent de la participante). À l'issue de cette période, le programme mettait automatiquement fin à l'essai en cours pour passer à l'essai suivant. Ces deux sessions étaient composées au total de 15 blocs de six essais, où l'ordre des phrases était aléatoire mais s'assurait que l'ordre de chaque essai dans un même bloc était différent. Chaque participante

a produit 270 énoncés au total, soit 90 répétitions de chaque phrase. Les six essais du premier bloc ont été considérés comme les essais de base car aucune altération du retour auditif n'était présente. Au cours de tous les essais des blocs suivants, la valeur du DAF a été choisie aléatoirement parmi 0, 60, 90, 120, 150 et 180 ms, et chaque valeur du DAF a été appliquée une fois dans chaque bloc. Il convient de noter que le DAF était présent de manière continue du début à la fin de la répétition d'une phrase.

Figure 9 Illustration des procédures de l'expérience 1



### 2.1.5 Traitement des données

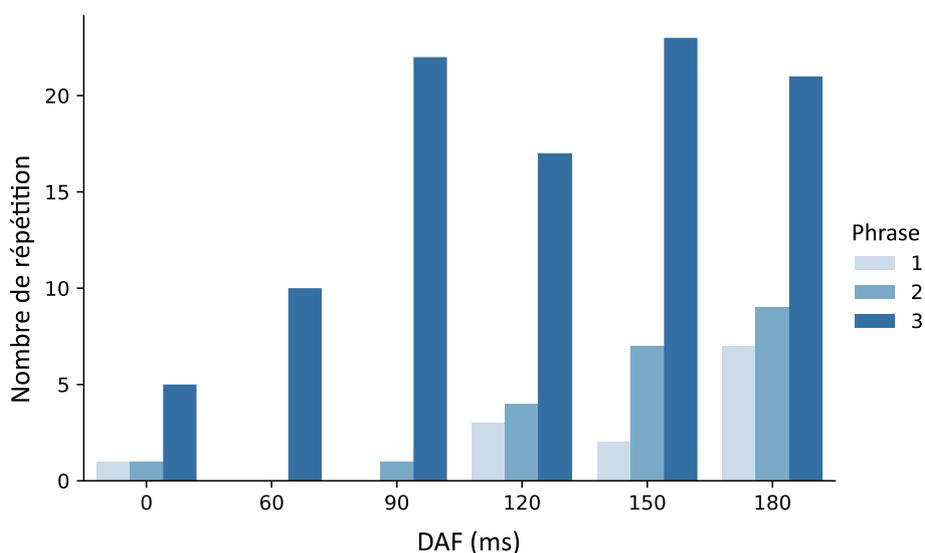
Toutes les analyses présentées dans cette thèse ont été basées sur des données acoustiques. Nous avons commencé par l'étiquetage manuel de toutes les frontières des phrases dans les enregistrements de chaque essai de chaque locutrice, ainsi que de toutes les frontières des voyelles au sein de chacune des répétitions des trois phrases dans les essais de base de chaque locutrice dans PRAAT (Boersma, 2001). Ensuite, les frontières des voyelles ont été projetées sur les autres répétitions des mêmes phrases prononcées par la même locutrice, en appliquant le *Dynamic Time*

*Warping* (Sakoe et Chiba, 1978) aux représentations cepstrales en Mel des énoncés, incluant les coefficients de 2 à 13 et leurs deltas. Les frontières ont ensuite été corrigées manuellement.

Les répétitions de la deuxième phrase ont été exclues des analyses en raison de la variation dans la production des troisième et cinquième syllabes par les participantes. Certaines d'entre elles ont prononcé des variantes telles que /man/ ou /manə/ au lieu de /mã/ et /nɔr mə/ au lieu de /nɔrm/. Cette variation entraîne une variabilité dans la durée des voyelles et dans l'organisation du rythme de la parole entre les locuteurs, ce qui peut affecter les résultats des analyses. En outre, les répétitions des deux autres phrases qui contenaient des disfluences et des erreurs ont été écartées des analyses.

Nous constatons une augmentation du nombre d'erreurs dans les répétitions avec l'augmentation du DAF. Parmi les trois phrases stimuli, la phrase 1 qui répète uniquement la consonne d'attaque des mots a généré le plus faible nombre d'erreurs. En revanche, la phrase 3 qui répète la consonne d'attaque des syllabes et certaines voyelles des noyaux suivants a engendré le plus grand nombre d'erreurs, comme l'illustre la Figure 10. De plus, la disparité du nombre d'erreurs entre la phrase 1 et la phrase 3 présente une tendance croissante en fonction du niveau du DAF.

Figure 10 Recensement du nombre de répétitions contenant des erreurs de production dans chaque type de phrase à différents niveaux du DAF dans l'expérience I



## 2.2 Expérience II Partie A – condition uniquement DAF

### 2.2.1 Objectives

S'appuyant sur les résultats de l'expérience I (cf. chapitre III [2.1.5](#) et chapitre IV [2.1](#)), nous avons ajusté nos paramètres expérimentaux pour une exploration plus approfondie de nos deux questions de recherche et des hypothèses (cf. chapitre III section 1). Ces ajustements seront explicités dans les sections « Matériaux » et « Procédures ».

Dans cette première phase de l'expérience II, nous avons uniquement appliqué le DAF pour examiner la première question de recherche, dont l'objectif principal est d'analyser la manière dont l'accentuation influence le contrôle sensorimoteur et la façon dont la coordination des composantes rythmiques s'adapte sous l'effet du DAF.

### 2.2.2 Participants

Dix locutrices, dont l'âge variait de 19 à 22 ans (âge moyen de 23,3 ans), et ayant le français comme langue maternelle, ont pris part à cette partie de l'expérience II. Toutes les participantes étaient des étudiantes de premier cycle à l'Université Sorbonne Nouvelle de Paris et ont reçu une compensation financière pour leur participation. Aucune d'entre elles n'a signalé de trouble de la parole ou de la voix, ni de déficience auditive. Les participantes ont signé un accord autorisant l'utilisation de leurs données à des fins d'analyse et de publication anonymes.

### 2.2.3 Matériaux

Les résultats de l'expérience I (cf. chapitre III [2.1.5](#)) ont révélé que certaines locutrices éprouvaient des difficultés à produire des virelangues sous l'effet du DAF. Dans le cadre de l'expérience II, les stimuli virelangues ont été substitués par trois phrases en français standard, présentant des mots avec des attaques syllabiques similaires, afin de préserver une régularité rythmique. Ces phrases avaient la même quantité de syllabes (cinq syllabes) et la même structure syntaxique (sujet-verbe-complément d'objet direct), mais différaient par leur complexité syllabique. La première phrase était principalement composée de syllabes de structure CV : /vivjẽ vi lə vẽ/ (« Vivien vit le vin. ») ; la deuxième de syllabes de structure CVC : /zaklin ʒɛʁ lə ʒuʁ/ (« Jacqueline gère le jour. ») ; et la troisième de syllabes de structure CCV(C) : /bʁadlɛ bʁiz lə bʁa/ (« Bradley brise

le bras. »). Nous ferons référence à ces phrases comme la phrase **CV**, la phrase **CVC** et la phrase **CCV(C)**.

Étant donné que les noyaux de syllabes plus complexes sont généralement plus courts, nous nous attendions à observer une variation systématique de la durée moyenne des voyelles d'une phrase à l'autre. Plus précisément, nous prévoyions que la durée moyenne des voyelles dans la phrase CV (c'est-à-dire celle ayant la structure syllabique la plus simple) serait plus longue que la durée moyenne des voyelles observée dans les deux autres types de phrases, même en tenant compte des différences de durée inhérentes à la qualité des voyelles (Maddieson, 1985 ; Rietveld et Frauenfelder, 1987). En utilisant ces phrases stimulantes, nous sommes en mesure d'opérer une variation systématique de la durée des voyelles d'origine de manière indépendante du niveau du DAF (voir chapitre IV [1.2](#)). Comme exposé dans la section [1](#) de ce chapitre, notre objectif en contrôlant cette variation de durée est de démêler l'effet de l'accentuation de celui de la durée.

La structure syntaxique joue un rôle significatif dans la délimitation des unités prosodiques (cf. chapitre I [2.2.1](#)). En utilisant des phrases présentant la même structure syntaxique comme stimuli, notre objectif était de minimiser les variations dans le phrasé des trois phrases, tant au sein des productions d'une même locutrice que parmi celles de plusieurs locutrices. Il est important de noter que dans les trois phrases de stimuli proposées, les niveaux de mot prosodique (Astésano, 2017) et de syntagme accentuel (Jun et Fougeron, 2000) peuvent se superposer (cf. chapitre I [2.2.1](#)). Par exemple, pour la phrase « Jacqueline gère le jour », trois mots prosodiques sont envisageables : « Jacqueline », « gère », et « le jour », avec les syllabes /lin/, /zɛʁ/, et /ʒuʁ/ portant l'accent principal de chaque unité. Ces trois unités peuvent également correspondre aux trois syntagmes accentuels. Dans nos analyses, nous désignerons ces unités sous le terme de « syntagme accentuel » (dorénavant SA). De plus, dans notre cas, chaque phrase dans son intégralité peut correspondre à un syntagme intonatif (dorénavant SI), avec la dernière syllabe (par exemple, /ʒuʁ/ dans /zaklin zɛʁ lə ʒuʁ/) comme tête de l'unité.

#### **2.2.4 Procédure**

Nous avons utilisé le même dispositif expérimental que celui utilisé dans l'expérience I cf. chapitre III [2.1.4](#)). Cependant, nous avons introduit de nouveaux stimuli (cf. chapitre III [2.2.3](#)) et avons apporté des modifications quant au nombre de blocs de six essais (16 blocs au total, un bloc de base y compris) et aux valeurs du DAF appliquées (0, 60 et 120 ms). Le niveau maximum du

DAF a été établi à 120 ms, en raison des constatations de l'expérience I qui indiquent que les effets du DAF atteignent un plateau autour de 120 ms (cf. chapitre IV [2.1](#)). Chaque niveau du DAF a été appliqué deux fois dans chaque bloc, et il y a eu un total de 96 répétitions de chaque phrase par chaque participante.

### **2.2.5 Traitement des données**

Les données acoustiques collectées ont été traitées de manière similaire à celles de l'expérience I (cf. chapitre III [2.1.5](#)), avec la différence que les frontières définies par la procédure du *Dynamic Time Warping* ont été utilisées directement pour les analyses. Cette méthode a été validée avec les données de l'expérience I en comparant les résultats obtenus à partir des données corrigées manuellement avec ceux obtenus à partir des données non corrigées. Comme pour l'expérience I, les répétitions de phrases contenant des disfluences ont été exclues des analyses.

## **2.3 Expérience II Partie B – condition perturbations multiples**

### **2.3.1 Objectives**

Dans cette seconde partie de l'expérience II, nous avons appliqué simultanément le DAF et un décalage continu de la  $f_0$  du retour auditif. De nouvelles participantes ont été recrutées spécifiquement pour cette partie de l'expérience, qui poursuit deux objectifs principaux.

Le premier objectif vise à examiner si une perturbation continue de la  $f_0$  du retour auditif induit un changement de la cible du contrôle sensorimoteur chez les locutrices, comme observé précédemment par Franken et al. (2021) (cf. chapitre II [3.3.2](#) et chapitre III [1.2](#)). Le deuxième objectif consiste à évaluer si la perturbation de la  $f_0$  du retour auditif entraîne une réduction de l'ampleur des réponses des locutrices au DAF, ce qui, selon notre hypothèse, correspond à une diminution du sens d'agentivité (cf. chapitre III [1.2](#)).

### **2.3.2 Participants, matériaux, procédure et traitement des données**

20 locutrices, dont l'âge variait de 18 à 25 ans (âge moyen de 21,5 ans), et ayant le français comme langue maternelle, ont pris part à cette partie de l'expérience II. Tout comme dans la

première partie de l'expérience, toutes les participantes étaient des étudiantes de premier cycle à l'Université Sorbonne Nouvelle de Paris et ont reçu une compensation financière pour leur participation. Aucune d'entre elles n'a signalé de trouble de la parole ou de la voix, ni de déficience auditive. Les participantes ont signé un accord autorisant l'utilisation de leurs données à des fins d'analyse et de publication anonymes.

Les mêmes stimuli ont été utilisés exactement comme dans la partie A de l'expérience (cf. [2.2.3](#)). Les 20 participantes ont été réparties en deux groupes, avec 10 participantes dans chaque groupe. Dans l'un des groupes, la  $f_0$  du retour auditif des locutrices a été positivement décalée (Groupe Positif), tandis que dans l'autre groupe, elle a été négativement décalée (Groupe Négatif).

Les procédures sont identiques à celles de la partie A (cf. [2.2.4](#)). Chaque participante a ainsi effectué 16 blocs de six répétitions (essais) des trois phrases de stimuli. Chaque niveau du DAF a été appliqué deux fois dans chaque bloc. La différence par rapport à la partie A réside dans le fait que dans les deux groupes de cette partie (Groupe Positif et Groupe Négatif), le décalage de la  $f_0$  a été appliqué au retour auditif des participantes. La valeur de ce décalage a été choisie de manière aléatoire pour chaque bloc parmi 0, 1 et 2 demi-tons (c'est-à-dire, la  $f_0$  n'a pas été décalée pour certains essais, tandis que pour d'autres, elle a été décalée de façon continue au sein de chaque bloc). Chaque valeur de décalage a été appliquée à cinq blocs.

Les données acoustiques collectées ont été traitées de la même manière que dans la partie A (cf. [2.2.5](#)).

### **3. Méthodes d'analyses**

Deux méthodes d'analyse, spécialement développées pour répondre de manière plus adéquate aux exigences d'investigation des questions de recherche formulées dans le cadre de cette thèse (cf. chapitre III section [1](#)), ont été proposées. La principale distinction entre ces deux méthodes réside dans l'objet d'étude. La première méthode d'analyse (section [3.1](#)), principalement utilisée pour étudier le rôle de l'accentuation dans la flexibilité du contrôle sensorimoteur, se fonde sur les unités prosodiques/phonologiques, tandis que la deuxième méthode (section [3.2](#)), employée pour examiner la coordination entre les composantes rythmiques, repose sur l'analyse de la fluctuation

de l'énergie dans le signal acoustique. Ainsi, la seconde méthode a une nature essentiellement physique.

Ces deux méthodes ont d'abord été appliquées aux données recueillies lors de l'expérience I (l'expérience pilote) afin de valider leur pertinence. Ces mêmes méthodes ont ensuite été appliquées aux données obtenues dans le cadre de l'expérience II, bien que certains paramètres aient été ajustés en fonction des résultats de l'expérience I et des objectifs de recherche. Par conséquent, dans les sections 3.1 et 3.2, nous exposerons également les différences dans la mise en œuvre des deux méthodes lors de l'analyse des données des deux expériences.

### **3.1 Mesurer la différence d'allongement entre les syllabes accentuées et les syllabes non accentuées sous l'effet du DAF**

Pour enquêter sur notre [première hypothèse](#), selon laquelle la flexibilité du contrôle sensorimoteur est influencée par l'accentuation à différents niveaux prosodiques, il est nécessaire d'évaluer la différence de l'effet d'allongement du DAF (chapitre II [1.2.1](#)) sur les syllabes accentuées et non accentuées. Cette évaluation s'est déroulée en deux étapes distinctes.

Dans la première étape, nous avons procédé à une comparaison directe de l'effet de l'allongement du DAF sur les voyelles<sup>1</sup> situées à différentes positions au sein de chaque phrase, en utilisant des modèles linéaires mixtes (cf. chapitre IV section [1](#)). L'objectif de cette première étape d'analyse était d'examiner de manière générale si l'effet du DAF variait en fonction de la position des voyelles.

Dans une seconde étape, nous avons quantifié la différence de durée entre les voyelles accentuées et non accentuées en calculant le ratio entre la durée des voyelles accentuées et des voyelles non accentuées. Par conséquent, si le DAF allonge davantage les voyelles accentuées, nous nous attendons à ce que ce ratio augmente sous l'effet du DAF. Dans les analyses préliminaires réalisées dans le cadre de l'expérience I (cf. chapitre III [2.1](#)), le ratio entre la durée totale des voyelles accentuables et des voyelles non accentuables a été calculé pour chaque répétition de phrase. Les voyelles accentuables sont définies comme les noyaux des syllabes finales des mots lexicaux qui sont susceptibles de porter l'accent d'un syntagme accentuel. Par exemple, dans la phrase 1 (/valmõ

---

<sup>1</sup> L'effet d'allongement du DAF est étudié en se penchant sur la durée des voyelles, car des recherches antérieures ont révélé que l'effet d'allongement du DAF se manifestait principalement au niveau des voyelles (Agnello, 1970).

vule vvar lə vəl/), les syllabes /mõ/, /lɛ/ /vwar/, /vəl/ et dans la phrase 3 (/mamã e mami mani ma mẽ/), les syllabes /mã/, /mi/, /ni/, /mẽ/ sont considérées comme des syllabes accentuables.<sup>2</sup>

Dans le cadre de la partie A « condition uniquement DAF » de l'expérience II (cf. chapitre III [2.2](#)), nous avons évalué plus précisément ce ratio au niveau du SA, puisque l'accentuation est associée aux unités prosodiques (cf. chapitre I [2.2.1](#)). Étant donné que les trois phrases de stimuli (cf. chapitre III [2.2.3](#)) présentaient la même structure syntaxique, nous nous attendions à une structure prosodique similaire pour les trois types de phrases dans les productions d'une même locutrice. Dans chaque répétition de phrase, **deux syntagmes accentuels** ont été pris en compte, à savoir le SA initial et le SA final. Par exemple, dans la phrase « Jacqueline gère le jour », « Jacqueline » a été considéré comme le SA initial et « le jour » comme le SA final. En réalité, comme discuté dans la section [2.2.3](#) de ce chapitre, dans certains cas, « gère le jour » pourrait également constituer un SA. Cependant, étant donné que notre analyse se concentre sur la variation de la différence de durée entre les voyelles accentuées et les voyelles non accentuées sous l'effet du DAF, le ratio entre la durée de « jour » et la durée de « le » est suffisamment représentatif pour rendre compte de cette variation.

Pour les dix locutrices qui ont participé à la partie A de l'expérience II (c'est-à-dire celles qui étaient dans la « condition uniquement DAF », cf. chapitre III [2.2](#)), nous avons vérifié si la deuxième syllabe (par exemple /lin/ dans le SA « Jacqueline ») de chaque phrase portait un ton montant ou un ton haut (Delais-Roussarie, 1993) afin de nous assurer que cette syllabe était accentuée, c'est-à-dire qu'elle marquait la frontière droite d'un SA. Parmi les dix locutrices, sept ont généralement formé leur SA initial pour chaque phrase avec les deux premières syllabes.

Ainsi, pour chaque répétition de phrase, deux types de ratios ont été calculés : 1) la durée de la deuxième voyelle d'une phrase (V2) divisée par la durée de la première voyelle (V1) ; 2) la durée de la dernière voyelle d'une phrase (V5) divisée par la durée de la quatrième voyelle (V4). D'un point de vue prosodique, V2 et V5, en tant que têtes de leurs unités, portent deux niveaux d'accentuation différents. Plus précisément, l'accentuation appliquée sur V2 délimite un SA, tandis que celle appliquée sur V5 délimite non seulement un SA, mais aussi un SI.

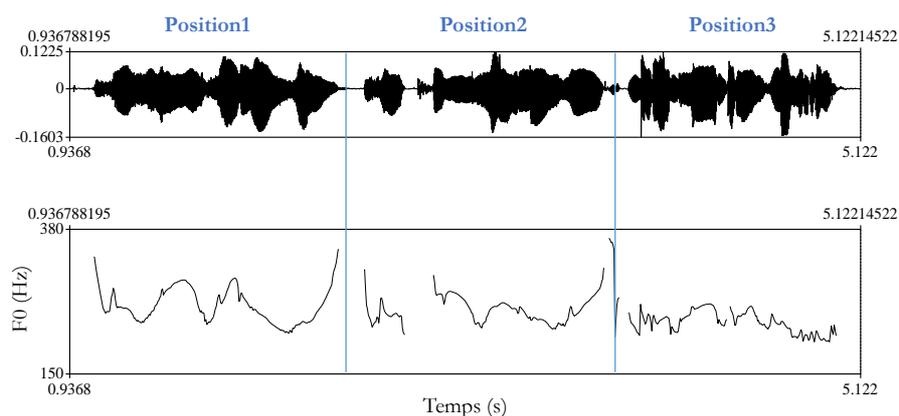
---

<sup>2</sup> La phrase 2 (/lə nɔrmã ɔr nɔrm mɔr/) a été exclue des analyses en raison de la variabilité observée dans les productions entre les participantes (cf. chapitre III [2.1.5](#)).

De plus, nous avons noté que la majorité des locutrices ont terminé les deux premières phrases dans un essai<sup>3</sup> (positions 1 et 2) avec un ton montant, tandis que la dernière phrase (position 3) était prononcée avec un ton descendant (cf. Figure 11). La courbe montante des phrases dans les positions 1 et 2 pourrait correspondre au phénomène de la continuation majeure (Delattre, 1966a). Cependant, la présence de la courbe montante pour les phrases en positions 1 et 2 n'était pas systématique. Par conséquent, nous avons divisé les données de ratio obtenues en **deux catégories d'analyse** : 1) les données des phrases prononcées en positions 1 et 2 (phrases non finales), produites par les locutrices qui affichaient simultanément V2 avec un ton haut/montant et V5 avec un ton montant dans ces positions (cinq locutrices) ; 2) les données des phrases prononcées en position finale de chaque essai, produites par les locutrices qui, dans cette position, affichaient V2 avec un ton haut/montant et V5 avec un ton descendant (sept locutrices).<sup>4</sup>

Figure 11 Exemple du signal acoustique et la courbe f0 d'un essai dans l'expérience II

Les lignes bleues verticales marquent les frontières des phrases.



<sup>3</sup> Un essai implique la répétition des trois phrases de stimuli (cf. chapitre III 2.2.4).

<sup>4</sup> Les catégories d'analyse ont été divisées en fonction de la position des phrases, plutôt que des SAs, afin de pouvoir comparer les effets du DAF entre les deux SAs d'une même phrase.

## 3.2 Mesurer la coordination entre la composante syllabique et la composante supra-syllabique

### 3.2.1 Présentation générale de la méthode

La méthode employée dans cette thèse pour évaluer la coordination entre la composante syllabique et la composante supra-syllabique repose sur l'analyse des modulations d'amplitude du signal acoustique de la parole. Ce signal est filtré en passe-bande afin de séparer les changements dus à la production des syllabes des changements plus lents résultant de la production des syllabes les plus proéminentes (cf. chapitre I 3.2.1). Ce processus génère deux signaux oscillatoires que nous nommerons syllAM (modulations d'amplitude syllabiques) et accAM (modulations d'amplitude des proéminences). Le premier signal devrait contenir des pics principalement au niveau des noyaux syllabiques (cf. chapitre I 1.4), tandis que les pics du second signal devraient être liés à la présence des syllabes plus proéminentes. Il est essentiel de noter que la proéminence, dans ce contexte, fait référence aux intervalles présentant une saillance d'énergie dans le signal de la parole, indépendamment d'autres paramètres (tels que la durée et la  $f_0$ ) qui peuvent être associés à la proéminence résultant de l'accentuation. Cette proéminence est donc un indice purement physique de bas niveau.

Les méthodes d'analyse basées sur la modulation d'amplitude des signaux acoustiques sont en constante évolution pour analyser les patrons rythmiques de la parole (par exemple, Tilsen et Arvaniti, 2013). Bien que ces signaux issus de la modulation d'amplitude ne tiennent pas compte d'autres dimensions de la parole, leurs caractéristiques se sont révélées être capables de discriminer entre les langues. Par exemple, Tilsen et Arvaniti (2013) ont démontré que la périodicité des proéminences supra-syllabiques est plus élevée en anglais qu'en allemand.

En raison de la nature oscillatoire des signaux obtenus, leur phase instantanée a été calculée, indiquant à chaque instant la position de chaque signal dans son cycle avec des valeurs angulaires allant de 0 à  $2\pi$ . Les valeurs de phase instantanée variant dans le temps ont permis de mesurer le degré de coordination (inversement liée à la variabilité de leur relation temporelle) entre syllAM et accAM dans différentes fenêtres temporelles en utilisant le calcul de la *Phase Locking Value* (PLV, cf. chapitre III 3.2.4). Cette mesure quantifie la variabilité du décalage entre les deux signaux (par rapport à la durée de leurs cycles) dans chaque fenêtre temporelle considérée.

Au cours des trois sections à venir, nous détaillerons le processus d'extraction des syllAM et accAM (3.2.2), la normalisation des signaux obtenus en vue de minimiser les artefacts (3.2.3), ainsi que la méthode de calcul de la PLV (3.2.4).

### 3.2.2 Extraction de la composante syllabique et la composante supra-syllabique

Le signal d'amplitude de chaque essai expérimental a été filtré en passe-bande pour extraire les signaux syllAM et accAM. Les fréquences de coupure des filtres utilisés pour extraire les signaux syllAM et accAM déterminent le degré de séparation entre les activités syllabique et supra-syllabique. Dans le contexte de cette thèse, il convient de noter que la parole enregistrée a été générée dans le cadre du DAF, ce qui signifie que le débit de parole peut présenter une grande variabilité. Par conséquent, nous avons adapté les fréquences de coupure des filtres à chaque enregistrement, c'est-à-dire à chaque essai expérimental. À l'exception de la fréquence de coupure basse de l'accAM fixée à 0,9 Hz, cette approche consistait à adapter les fréquences de coupure de chaque enregistrement analysé selon :

$$sLow = \frac{(syllR + accR)}{2}$$
$$sHi = syllR + (syllR - sLow)$$

où *sLow* indique à la fois la fréquence de coupure basse de syllAM et la fréquence de coupure haute de accAM ; *sHi* indique la fréquence de coupure haute de syllAM ; *syllR* et *accR* représentent respectivement les débits de syllabes et d'accents attendus. Ceux-ci ont été obtenus en divisant le nombre moyen de syllabes et d'accents attendus des trois phrases dans un essai par leur durée moyenne.

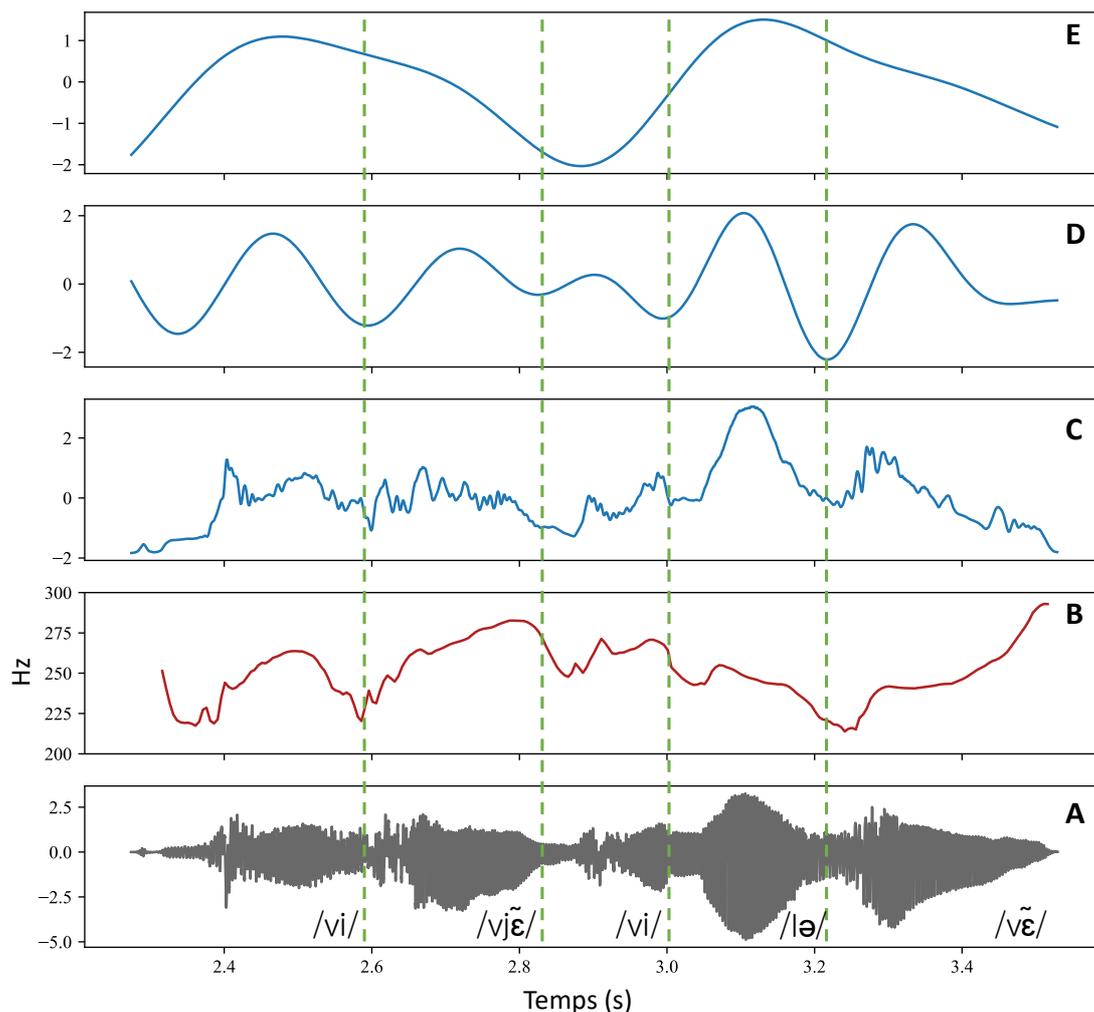
La Figure 12 présente un exemple des deux signaux (syllAM et accAM) extraits à travers l'analyse des modulations d'amplitude de la phrase « Vivien vit le vin » (/vivjẽ vi lə vẽ/). Nous pouvons observer que les transitions entre les cycles dans le syllAM (panneau D) sont étroitement corrélées aux frontières des syllabes, marquées par les lignes horizontales en vert. Les deux voyelles phonologiquement accentuées /vjẽ/ et /vẽ/ se distinguent par une courbe ascendante de la f0 (panneau B). Dans cet exemple particulier, les pics supra-syllabiques (dans le panneau E) ne coïncident pas avec les syllabes accentuées. Il est important de noter que l'intensité est un indice moins significatif de l'accentuation en français (Delattre, 1966b). Les pics supra-syllabiques illustrés

dans cet exemple correspondent aux points saillants de l'énergie physique dans la production, indépendamment de l'accentuation.

Figure 12 Modulations d'amplitude pour la phrase « Vivien vit le vin » (/vivjẽ vi lə vẽ/)

A : le signal acoustique ; B : la courbe de la  $f_0$  ; C : la variation de l'intensité ; D : l'oscillation syllabique ; E : l'oscillation supra-syllabique

Dans A, C, D et E, les valeurs Y correspondent aux z-scores. Les lignes verticales en vert marquent les frontières des syllabes.



### 3.2.3 Normalisation d'amplitude

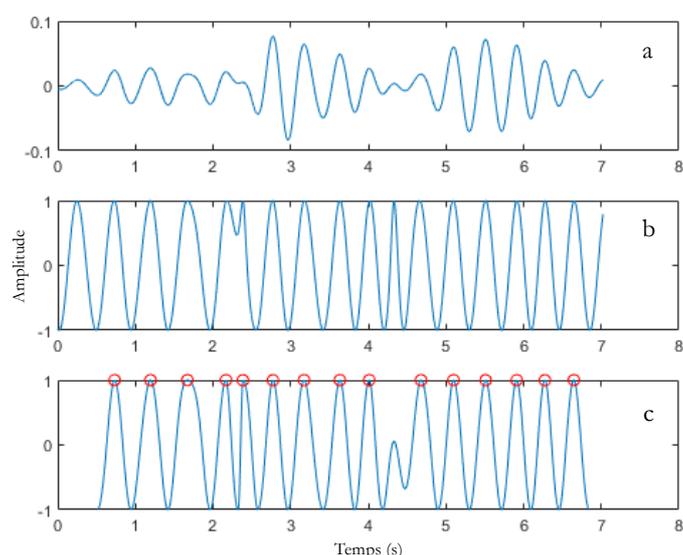
La procédure d'estimation des valeurs de phase instantanées des signaux oscillatoires repose généralement sur l'application de la transformée de Hilbert, qui suppose que la phase varie beaucoup plus rapidement que l'amplitude (Huang et al., 2009). Cependant, cette hypothèse n'est pas applicable dans le cas des signaux AM, où les amplitudes des cycles peuvent différer considérablement entre cycles consécutifs. Ainsi, pour obtenir la phase instantanée des signaux

syllAM et accAM, il est possible de normaliser l'amplitude des signaux en amenant les valeurs minimales et maximales locales à des valeurs constantes de -1 et 1, respectivement.

Il est important de noter que les pics parasites peuvent fausser l'estimation de la phase, en particulier lorsque plusieurs pics se produisent dans un seul cycle de la série temporelle. Cette situation peut se produire, par exemple, lorsque la production de segments entraîne des pics supplémentaires dans le signal syllAM. Une normalisation aveugle (le panneau b dans la Figure 13) peut alors introduire des cycles parasites et affecter l'estimation de la phase.

Figure 13 Procédure de normalisation

Panneau a : signal syllAM. Panneau b : signal syllAM normalisé basée sur tous les maxima. Panneau c : signal syllAM normalisé basée sur un maximum par voyelle. Les cercles rouges indiquent les maxima sélectionnés.



Afin de résoudre ce problème, nous avons adopté une stratégie qui consistait à sélectionner un seul pic par voyelle (le panneau c dans la Figure 13), sachant que l'emplacement de chaque voyelle était connu pour les données analysées (cf. chapitre III 2.1.5 sur le traitement des données). Si un signal syllAM correspondant à une phrase présentait plus de pics que de voyelles attendues, nous avons éliminé les pics en excès en les associant au centre vocalique le plus proche (défini comme le point médian entre le début et la fin) et en conservant pour chaque voyelle le pic le plus saillant. Les vallées ont été déterminées en recherchant la plus petite valeur de syllAM apparaissant entre deux pics consécutifs. La même procédure a été appliquée au signal accAM en conservant uniquement le nombre de pics n'excédant pas le nombre de syllabes pouvant être accentuées dans chaque phrase. Ensuite, les valeurs de chaque intervalle commençant par un minimum local et se terminant par le maximum local suivant ont été normalisées en utilisant la formule suivante :

$$\hat{X}_{i+j} = 2 \times \left( \frac{X_{i+j} - X_i}{|X_{i+k} - X_i| - \text{sign}(X_{i+k} - X_i)} \right)$$

où  $X_i$  est la valeur de  $X$  au  $i^{\text{ème}}$  minimum local,  $X_{i+j}$  est la valeur de  $X$  observée  $j$  échantillons après  $X_i$ , et  $k$  est le nombre d'échantillons séparant  $X_i$  du maximum local suivant (de sorte que  $X_{i+k}$  est la valeur de  $X$  au maximum local suivant  $X_i$ ). Enfin,  $|\cdot|$  indique la valeur absolue et  $\text{sign}(\cdot)$  indique la fonction de *signe*.

### 3.2.4 Calcul de l'indice de coordination – Phase Locking Value (PLV)

L'avancement relatif de deux signaux  $X$  et  $Y$  (ex. syllAM et accAM) dans leurs cycles à l'instant  $t$  peut être exprimé par la différence de phase généralisée :

$$\Phi_{X,Y}(t) = m \times \Phi_X(t) - n \times \Phi_Y(t)$$

où  $m$  et  $n$  sont deux entiers déterminant le ratio entre les fréquences ( $\omega_X$  et  $\omega_Y$ ) de  $X$  et  $Y$  (de sorte que  $n \times \omega_X = m \times \omega_Y$ ) et  $\Phi_X(t)$  et  $\Phi_Y(t)$  sont les valeurs de phase à l'instant  $t$ .

La méthode d'extraction de la phase instantanée à partir des courbes telles que celle présentée dans la Figure 13 peut présenter des artefacts dus à la présence des pics parasites, tels que le pic entre 4 et 5 secondes dont l'amplitude reste faible après la normalisation. Afin d'éliminer ces artefacts, nous avons mis en place une méthode en deux étapes. Tout d'abord, nous avons débruité les valeurs de phase obtenues en appliquant la transformée de Hilbert à la série temporelle d'amplitude normalisée. Ensuite, nous avons exclu de l'analyse les valeurs de différence de phase généralisée correspondant aux points où l'une ou l'autre des valeurs de phase débruitée diminuait.

La mesure de la coordination entre les deux signaux pendant un intervalle de temps peut être évaluée en utilisant l'intensité du premier moment de Fourier de la distribution des valeurs de différence de phase généralisée dans cet intervalle (Rosenblum et al., 2001).

$$PLV = \langle \cos \Phi_{X,Y}(t) \rangle^2 + \langle \sin \Phi_{X,Y}(t) \rangle^2$$

où les crochets indiquent la moyenne. Le coefficient de coordination, appelé PLV, est un indice de la force de la coordination entre les deux signaux, qui varie de 0 (pour une distribution uniforme) à 1 (pour une distribution de valeurs identiques), et qui est inversement lié au degré de variabilité de la différence de phase généralisée dans l'intervalle de temps considéré. Nous avons calculé une PLV pour chaque répétition de chaque phrase. Étant donné que cet indice dépend de la largeur de la distribution des valeurs de la différence de phase généralisée, il est influencé par la relation entre le nombre de cycles observés dans une fenêtre d'analyse et la fréquence d'échantillonnage des

valeurs de phase. Afin de minimiser cette influence, nous avons interpolé les valeurs  $\Phi_{X,Y}$  pour obtenir un nombre constant d'observations (1000) dans chaque fenêtre temporelle égale à la durée d'une phrase, ce qui nous permet d'obtenir des comparaisons significatives entre les répétitions (car le nombre de pics devrait dépendre du nombre de syllabes et d'accents qui restent constants entre les répétitions des mêmes phrases).

## **4. Résumé du chapitre**

Dans ce chapitre, nous avons tout d'abord exposé nos quatre hypothèses en relation avec nos deux questions de recherche. En substance, nous postulons une interaction entre le rôle organisateur de l'accentuation et celui du sens d'agentivité dans les processus de contrôle sensorimoteur, ainsi qu'une adaptation de l'organisation rythmique (la coordination entre les composantes rythmique) au sein de cette interaction.

Nous avons ensuite présenté les deux expériences portant sur les perturbations du retour auditif (retard du retour auditif et/ou décalage continu de la  $f_0$  du retour auditif) que nous avons menées pour examiner nos hypothèses. En analysant la manière dont l'accentuation et le sens d'agentivité influencent l'adaptation de la production de la parole des participantes dans diverses conditions de perturbations du retour auditif, nous avons pu évaluer leur rôle dans le contrôle sensorimoteur.

Enfin, nous avons introduit les deux méthodes que nous avons utilisées pour analyser nos données. La première méthode, basée sur la durée des syllabes, avait pour objectif principal d'évaluer le rôle de l'accentuation dans le contrôle sensorimoteur dans le contexte du DAF. La deuxième méthode, basée sur l'analyse des modulations d'amplitude, visait à étudier l'organisation rythmique, indépendamment des analyses des unités phonologiques (par exemple, la syllabe, le syntagme accentuel).

Dans les trois prochains chapitres (IV, V, VI), nous présenterons les analyses fondées sur ces deux méthodes. Les analyses exposées dans chaque chapitre seront destinées à répondre à une ou deux hypothèses spécifiques.

---

# CHAPITRE IV : ANALYSE DU RÔLE DE L'ACCENTUATION DANS LE CONTRÔLE SENSORIMOTEUR PENDANT LA PRODUCTION DE LA PAROLE

## Sommaire du chapitre

<b>1. Comparer l'effet d'allongement du DAF sur la durée des voyelles dans différentes positions</b> .....	<b>78</b>
1.1 Résultats de l'expérience I .....	78
1.2 Résultats de l'expérience II (« condition uniquement DAF ») .....	80
1.3 Conclusion.....	82
<b>2. Analyse du rôle organisateur de l'accentuation</b> .....	<b>83</b>
2.1 Effet du DAF sur la différence de durée entre les voyelles accentuables et non accentuables – expérience I.....	83
2.2 Effet du DAF sur la différence de durée entre les voyelles accentuées et non accentuées – expérience II (« condition uniquement DAF ») .....	85
2.3 Effets d'allongement du DAF interagissant avec l'allongement final .....	87
<b>3. Discussion du chapitre</b> .....	<b>88</b>

Dans ce chapitre, nous présenterons les résultats des analyses visant à explorer notre [première hypothèse](#) (La flexibilité du contrôle sensorimoteur est influencée par l'accentuation à différents niveaux prosodiques). Pour chaque type d'analyse (cf. chapitre III [3.1](#)), nous examinerons et comparerons les résultats de l'expérience I (cf. chapitre III [2.1](#)) et de l'expérience II (« condition uniquement DAF », cf. chapitre III [2.2](#)), le cas échéant.

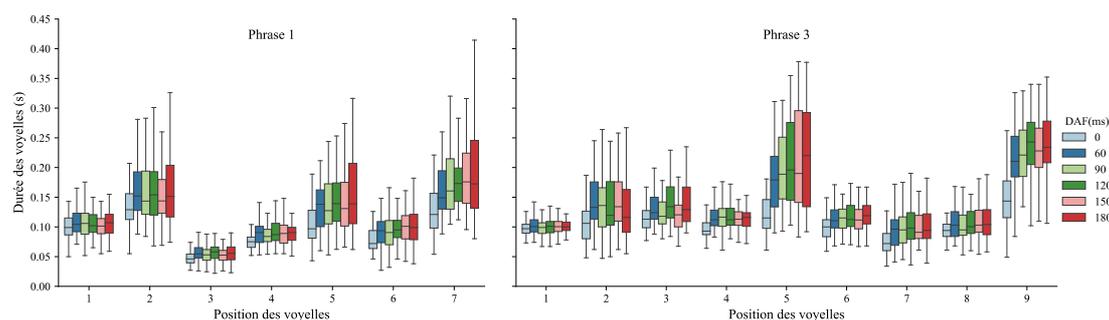
Premièrement, nous comparerons si l'effet d'allongement du DAF sur la durée des voyelles varie en fonction de leur position dans les énoncés (section [1](#)). Ensuite, pour évaluer le rôle de l'accentuation dans le contrôle sensorimoteur lors de la production de la parole, nous analyserons l'effet du DAF sur la différence de durée entre les voyelles accentuées et non accentuées (section [2](#)).

## 1. Comparer l'effet d'allongement du DAF sur la durée des voyelles dans différentes positions

### 1.1 Résultats de l'expérience I

La Figure 14 présente la durée des voyelles produites à différents niveaux du DAF (0, 60, 90, 120, 150 et 180 ms, différenciés par couleur) dans différentes positions au sein des phrases 1 (/valmõ vule vvar læ vøl/) et 3 (/mamã e mami mani ma mẽ/). La phrase 2 a été exclue des analyses (cf. chapitre III 2.1.5). En examinant la Figure 14, nous pouvons observer que la durée des voyelles dans chaque position des énoncés tend à augmenter sous l'influence du DAF. De plus, il semble que le DAF provoque un allongement plus prononcé des voyelles dans certaines positions.

Figure 14 Durée des voyelles dans différentes positions, respectivement pour la phrase 1 (/valmõ vule vvar læ vøl/) et la phrase 3 (/mamã e mami mani ma mẽ/), dans l'expérience I



Nous avons conduit *deux modèles linéaires mixtes*, séparément pour la phrase 1 et la phrase 3, afin d'examiner l'effet du DAF sur la durée des voyelles dans différentes positions. Les prédicteurs étaient le niveau du DAF (référence : 0 ms), la position des voyelles, l'ordre de production de la phrase dans un essai (référence : position 1) et leur double interaction<sup>5</sup>. L'effet de la position des voyelles a été codé par un contraste basé sur la déviation de la valeur moyenne, ce qui a donné lieu à un intercept correspondant au comportement moyen observé sur les voyelles en toutes les positions. Les effets aléatoires comprenaient un intercept par sujet pour chaque variable prédictive<sup>6</sup>.

<sup>5</sup> Nous avons essayé d'implémenter toutes les interactions possibles (doubles et triples) entre les prédicteurs dans tous les modèles dans cette thèse. Dans certains modèles, une ou plusieurs interactions n'étaient pas significatives et ne contribuaient pas à l'adéquation du modèle (évalué par un test du ratio de vraisemblance entre le modèle avec l'interaction et le modèle sans interaction à l'aide du paquet MASS [Venables et Ripley, 2002]). Ces interactions ont donc été exclues et ne sont pas mentionnées dans les descriptions des modèles.

<sup>6</sup> La structure des effets aléatoires contenant un intercept aléatoire spécifique au locuteur pour chaque prédicteur du modèle a été appliquée à tous les modèles décrits dans cette thèse sauf quelques exceptions dans le chapitre IX.

Les données des essais de base n'ont pas été incluses dans le modèle. Le modèle pour la phrase 1 comportait un total de 3402 observations, tandis que celui pour la phrase 3 comprenait un total de 3534 observations (les résultats clé des modèles sont reportés dans les Tableaux [1](#) et [2](#) dans l'Annexe).

### **Résultats de la phrase 1 :**

Dans les répétitions de la phrase 1 en position 1, la durée des V2 et V7 (c'est-à-dire les voyelles dans les syllabes /mɔ̃/ et /vɔl/) est plus grande que la durée moyenne des voyelles tandis que la durée de V3 (c'est-à-dire la voyelle dans la syllabe /vu/) est plus petite que la durée moyenne. La durée de V5 (c'est-à-dire la voyelle dans la syllabe /vwar/) est plus grande dans les répétitions en positions 2 et 3 ainsi que la durée de V2 dans les répétitions en position 3. La durée de V7 est plus petite dans les répétitions en position 3.

En moyenne, tous les niveaux du DAF allongent significativement les voyelles. En général, l'effet d'allongement du DAF à 120 et 180 ms est plus forte dans les répétitions en position 2. L'effet d'allongement du DAF varie en fonction de la position des voyelles. L'effet de tous les niveaux du DAF est plus fort sur les V2, V5 et V7 que l'effet moyen de toutes les voyelles. En revanche, il est moins fort sur la V4 (c'est-à-dire la voyelle dans la syllabe /lɛ/). L'effet du DAF à 180 ms est également moins fort sur la V6 (c'est-à-dire la voyelle dans la syllabe /lɔ̃/).

### **Résultats de la phrase 2 :**

La durée des voyelles dans les répétitions de la phrase 3 en position 3 est en moyenne inférieure à celle des répétitions en position 1. Dans les répétitions en position 1, nous observons que la durée de V7 (la voyelle dans la syllabe /ni/) est inférieure à la durée moyenne des voyelles, tandis que la durée de V9 (la voyelle dans la syllabe /mɛ̃/) est supérieure à la moyenne. Quant à la durée de V2 (la voyelle dans la syllabe /mã/), elle ne présente pas de différence significative par rapport à la durée moyenne des voyelles, bien que nous observions une légère tendance à être plus courte que la durée moyenne. Toutefois, dans les répétitions en position 3, cette tendance est atténuée. Nous remarquons également que la différence entre la durée de V9 et la durée moyenne des voyelles est réduite en position 2 et 3.

---

Nous ne mentionnerons plus cette description des effets aléatoires dans le reste de cette thèse si les exceptions ne sont pas concernées.

Tout comme pour la phrase 1, l'ensemble des niveaux du DAF ont significativement augmenté la durée des voyelles. L'effet du DAF à 60 ms est encore plus prononcé dans les répétitions en position 2 et 3. De même, l'effet du DAF à 150 ms est plus marqué dans les répétitions en position 3.

L'ampleur de l'effet du DAF varie en fonction de la position des voyelles et de la position des répétitions de la phrase. Nous observons que le DAF à 90 ms a un impact plus important sur la durée de V2 que la durée moyenne des voyelles. De même, l'effet du DAF à 180 ms a tendance à être plus fort sur la durée de V2 (à l'exception la V2 dans les répétitions en position 3). De manière générale, l'effet de tous les niveaux du DAF est plus marqué sur la durée de V5 (la voyelle dans la syllabe /mi/) que sur la durée moyenne des voyelles. Par ailleurs, l'effet de tous les niveaux du DAF a également tendance à être plus important sur la durée de V9 que sur la durée moyenne des voyelles, surtout dans les répétitions en positions 2 et 3. En revanche, l'effet du DAF à 90 ms et 150 ms est moins prononcé sur la durée de V3 (/e/) que sur la durée moyenne des voyelles. De même, l'effet du DAF à 180 ms est moins important sur la durée de V4 (la voyelle dans la deuxième syllabe /ma/). Enfin, nous constatons que l'effet de tous les niveaux du DAF (à l'exception du DAF à 60 ms) est significativement moins important sur la durée de V6 (la voyelle dans la troisième syllabe /ma/).

### **Conclusion :**

L'effet de l'allongement du DAF n'est pas uniforme sur toutes les voyelles et varie selon leur position dans la phrase.

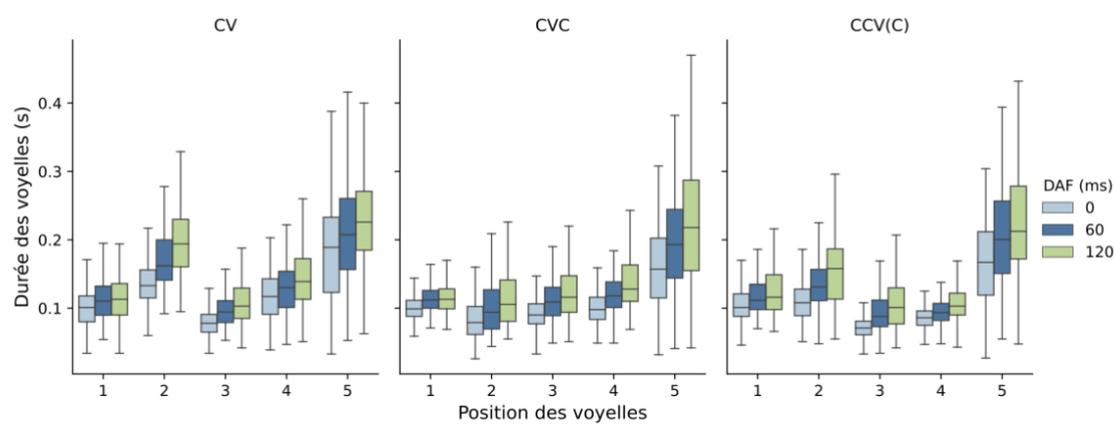
## **1.2 Résultats de l'expérience II (« condition uniquement DAF »)**

La Figure 15 illustre la durée des voyelles dans différentes positions au sein de chaque phrase, à différents niveaux du DAF (0, 60, 120 ms, différenciés par couleur). Nous avons conduit *un modèle linéaire mixte* pour déterminer si le DAF allongeait les voyelles de manière différente selon les positions ou les phrases. Les prédicteurs du modèle comprenaient le niveau du DAF (référence : 0 ms), la position de la voyelle, le type de phrase et leurs doubles et triples interactions. Les effets de la position de la voyelle et du type de phrase ont été codés en termes d'écart par rapport au contraste

moyen. Nous avons exclu les données des essais de base du modèle, qui contient un total de 13025 observations (les résultats clé des modèles sont reportés dans le Tableau 3 dans l'Annexe).

Les résultats seront résumés en deux parties. Tout d'abord, nous résumerons l'effet de la structure syllabique sur la durée des voyelles et la comparaison de l'effet du DAF sur différents types de phrases. Ensuite, nous présenterons un résumé de l'effet d'allongement du DAF sur les voyelles dans différentes positions.

Figure 15 Durée des voyelles dans différentes positions, respectivement pour chaque type de phrase (CV, CVC, CCV(C)), dans l'expérience II (« condition uniquement DAF »)



### **Résultats : effet d'allongement du DAF non prédit par la durée des syllabes différent en structure**

Dans les essais sans DAF, la durée des voyelles dans la phrase CV est significativement supérieure à la durée moyenne des voyelles calculée pour les trois types de phrases. La différence entre la durée de la deuxième et de la cinquième voyelle dans la phrase CV et la durée moyenne de ces mêmes voyelles calculée pour tous les types de phrases est encore plus importante. Cependant, ce n'est pas le cas pour la troisième voyelle de la phrase CV, dont la durée ne diffère pas significativement de la durée moyenne de la troisième voyelle calculée sur l'ensemble des types de phrases.

En revanche, la durée des voyelles dans la phrase CCV(C) est significativement inférieure à la durée moyenne des voyelles pour l'ensemble des types de phrases. L'écart entre la durée de la troisième et de la quatrième voyelle dans la phrase CCV(C) et la durée moyenne des mêmes voyelles observée dans différents types de phrases présente une différence encore plus significative. Néanmoins, ce n'est pas le cas pour la deuxième voyelle de la phrase CCV(C), dont la durée tend à moins différer de la durée moyenne de la deuxième voyelle calculée à travers les types de phrases.

En ce qui concerne les effets du DAF, les deux niveaux du DAF (soit 60 ms et 120 ms) entraînent en général une augmentation de la durée des voyelles. Malgré la durée relativement longue des voyelles dans la phrase CV et la courte durée des voyelles dans la phrase CCV(C), aucune variation significative de l'effet du DAF sur la durée des voyelles n'est observée à travers les trois types de phrases. En d'autres termes, les différences de durée des voyelles dues à la structure syllabique ne modulent pas de manière significative l'allongement dû au DAF.

### **Résultats : effet d'allongement du DAF sur la durée des voyelles dans différentes positions**

L'ampleur de l'effet du DAF varie en fonction de la position de la voyelle. En particulier, l'effet d'allongement dû au DAF est plus fort sur les voyelles en deuxième et cinquième position que l'effet moyen sur l'ensemble des voyelles dans les trois types de phrases, tandis que l'effet sur les voyelles en troisième et quatrième position est plus faible que l'effet moyen.

En outre, l'effet du DAF est plus fort sur la deuxième voyelle de la phrase CV par rapport à l'effet moyen sur les trois types de phrases. L'effet du DAF à 120 ms sur la quatrième voyelle des phrases CV et CCV(C) est plus faible que l'effet moyen pour les trois types de phrases. L'effet du DAF à 60 ms sur la cinquième voyelle de la phrase CV est également plus faible que l'effet moyen dans les trois types de phrases, tandis que l'effet du DAF à 60 ms sur la cinquième voyelle de la phrase CCV(C) est plus fort que l'effet moyen dans les trois types de phrases.

En somme, l'effet de l'allongement du DAF n'est pas uniforme sur toutes les voyelles et varie selon leur position dans la phrase.

### **1.3 Conclusion**

En résumé, les résultats des deux expériences sont cohérents. Plus précisément, différentes parties de l'énoncé, c'est-à-dire les voyelles dans différentes positions de la phrase, ne présentent pas le même niveau d'adaptation sous l'effet du DAF. De plus, les résultats de l'expérience II nous confirment que ce niveau d'adaptation ne peut pas être prédit par la durée des voyelles.

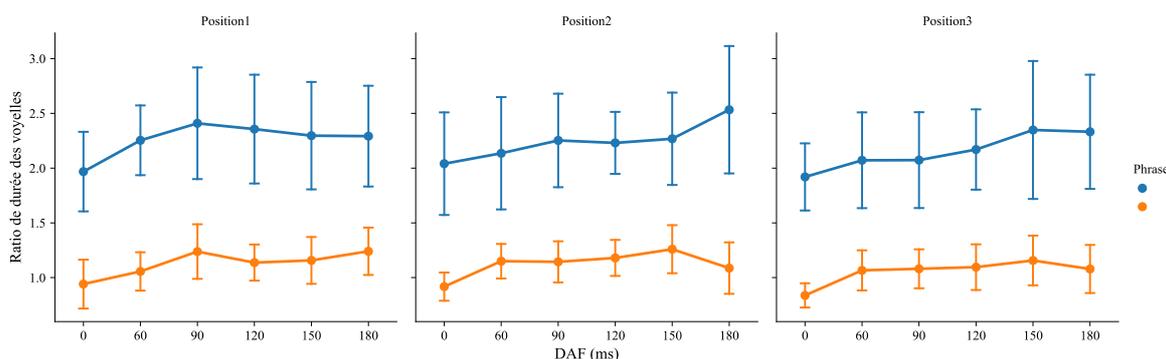
## 2. Analyse du rôle organisateur de l'accentuation

Pour analyser l'effet du DAF sur la différence de durée entre les voyelles accentuées et non accentuées, nous avons étudié l'effet du DAF sur le ratio entre la durée des voyelles accentuées et non accentuées (cf. chapitre III 3.1). Nous avons d'abord effectué une analyse préliminaire avec les données de l'expérience I (section 2.1), et avons ensuite affiné nos analyses en utilisant les données de l'expérience II, en prenant en compte les unités prosodiques et les différents niveaux d'accentuation (sections 2.2 et 2.3).

### 2.1 Effet du DAF sur la différence de durée entre les voyelles accentuables et non accentuables – expérience I

Dans le cadre de l'expérience I, nous avons analysé l'effet du DAF sur le ratio entre la durée totale des voyelles accentuables et non accentuables (ratioDur) (cf. chapitre III 3.1). La Figure 16 présente la moyenne et l'écart-type du ratioDur à chaque niveau du DAF (0, 60, 90, 120, 150 ou 180 ms) pour les phrases 1 (courbes bleues) et 3 (courbes oranges) dans chaque position des essais expérimentaux.

Figure 16 Moyenne et écart-type du ratio entre la durée totale des voyelles accentuables et non accentuables pour la phrase 1 (/valmɔ̃ vuʁ vvar lə vɔl/) et la phrase 3 (/mamã e mami mani ma mɛ̃/) dans chaque position des essais expérimentaux dans l'expérience I



Un modèle linéaire mixte a été conduit afin d'évaluer l'effet du DAF sur le ratioDur. Les prédicteurs incluaient le niveau du DAF (avec un codage basé sur un contraste par différences

successives, testant les différences entre les valeurs successifs : 0-60-90-120-150-180 ms<sup>7</sup>), la phrase (référence : phrase 1), l'ordre de production de la phrase dans un essai (référence : position 1), ainsi que leurs doubles et triples interactions. Les données correspondant aux essais de base ont été exclues. Le modèle a inclus un total de 881 observations (les résultats clé des modèles sont reportés dans le Tableau 4 dans l'Annexe).

### **Résumé des résultats :**

La phrase 3 (cf. les courbes en orange dans la Figure 16), présente une proportion moins élevée de syllabes accentuables par rapport à la phrase 1 (cf. les courbes en bleu dans la Figure 16)<sup>8</sup>, ce qui se traduit naturellement par un ratioDur plus faible. Pour la phrase 1, le ratioDur est plus faible en position 3 qu'en position 1.

Lorsque nous examinons la phrase 1 en position 1, les essais avec un DAF à 60 ms ont montré une augmentation significative du ratioDur par rapport aux essais sans DAF. Le ratioDur continue à augmenter dans les essais avec 90 ms du DAF et l'effet du DAF à 120 ms n'est pas significativement différent de celui observé à 90 ms. Par ailleurs, l'ampleur des effets du DAF entre deux niveaux successifs de plus de 120 ms (c'est-à-dire entre 120 ms et 150 ms, entre 150 ms et 180 ms) n'est pas significativement différente. Un patron similaire d'effet du DAF est observé pour la phrase 1 en position 3 et pour la phrase 3 dans toutes les positions.

Dans la phrase 1 en position 2, l'effet du DAF à 180 ms est plus important que celui observé à 150 ms. En revanche, cet effet n'est pas présent dans la phrase 3 en position 2.

### **Conclusion :**

Le DAF a tendance à allonger davantage la durée des voyelles dans les syllabes accentuables par rapport à celles dans les syllabes non accentuables. En d'autres termes, la différence de durée entre les deux types de voyelles augmente avec l'augmentation du niveau du DAF. Cependant, cet effet atteint un plateau vers 120 ms de DAF. De manière générale, cet effet du DAF ne présente pas de différence significative entre la phrase difficile à produire (la phrase 3) et la phrase relativement moins difficile à produire (la phrase 1) (cf. chapitre III 2.1.5), ce qui suggère que cet effet du DAF est indépendant de la difficulté de la tâche de production.

---

<sup>7</sup> Cette option permet de déterminer le délai (c'est-à-dire le niveau du DAF) à partir duquel l'allongement causé par la perturbation cesse d'augmenter.

<sup>8</sup> La phrase 1 (/valmɔ̃ vule vvar lə vɔl/) contient sept syllabes, dont quatre sont accentuables (/mɔ̃/, /le/, /vvar/, /vɔl/). La phrase 3 (/mamã e mami mani ma mɛ̃/) contient neuf syllabes, dont quatre sont accentuables (/mã/, /mi/, /ni/, /mɛ̃/).

## 2.2 Effet du DAF sur la différence de durée entre les voyelles accentuées et non accentuées – expérience II (« condition uniquement DAF »)

Dans le cadre de la « condition uniquement DAF » de l'expérience II (cf. chapitre III [2.2](#)), l'effet du DAF a été analysé en fonction de deux types de ratio : [1] le ratio entre la durée de la voyelle accentuée V2 et celle de la voyelle non accentuée V1 dans le SA initial de chaque phrase ; [2] le ratio entre la durée de la voyelle accentuée V5 et celle de la voyelle non accentuée V4 dans le SA final de chaque phrase. De plus, les données ont été réparties en deux catégories d'analyse : [1] les phrases non finales en positions 1 et 2 des essais expérimentaux ; [2] les phrases finales en position 3 des essais expérimentaux (voir la section [3.1](#) du chapitre III pour plus de détail).

La Figure [17](#) affiche la moyenne et l'écart-type des deux types de ratios (différenciés par couleur) à chaque niveau du DAF (0, 60 ou 120 ms), pour chaque catégorie d'analyse (catégorie 1 en haut et catégorie 2 en bas). *Deux modèles linéaires mixtes* ont été conduits pour chaque catégorie d'analyse en utilisant des prédicteurs comprenant le niveau du DAF (référence : 0 ms), le type de ratio (référence : V2/V1), le type de phrase, ainsi que leurs doubles interactions. L'effet du type de phrase a été codé en utilisant un écart par rapport au contraste moyen. Les données des essais de base n'ont pas été incluses dans les modèles. Le modèle pour la catégorie d'analyse 1 comportait un total de 1776 observations, tandis que celui pour la catégorie d'analyse 2 comprenait un total de 1170 observations (les résultats clé des modèles sont reportés dans le Tableau [5](#) dans l'Annexe).

### **Résultats de la catégorie d'analyse 1 (phrases non finales) :**

Les résultats de la catégorie d'analyse 1 indiquent que, en l'absence du DAF, le ratio V5/V4 est généralement plus élevé que le ratio V2/V1 pour toutes les phrases, bien que l'écart entre les deux ratios soit plus faible dans la phrase CV et plus important dans la phrase CCV(C) par rapport à l'écart moyen des trois types de phrases.

Lorsque le DAF est présent, celui-ci a tendance à augmenter le ratio V2/V1 pour tous les types de phrases, mais cette augmentation n'est significative qu'au niveau plus élevé du DAF (120 ms). La ratio V2/V1 dans la phrase CCV(C) est moins affectée par le DAF à 120 ms. En ce qui concerne le ratio V5/V4, le DAF à 60 ms a tendance à augmenter ce ratio plus que le ratio V2/V1 (cependant, cette différence n'est que presque significative). Cependant, le DAF à 120 ms a un effet plus faible sur le ratio V5/V4 que sur le ratio V2/V1.

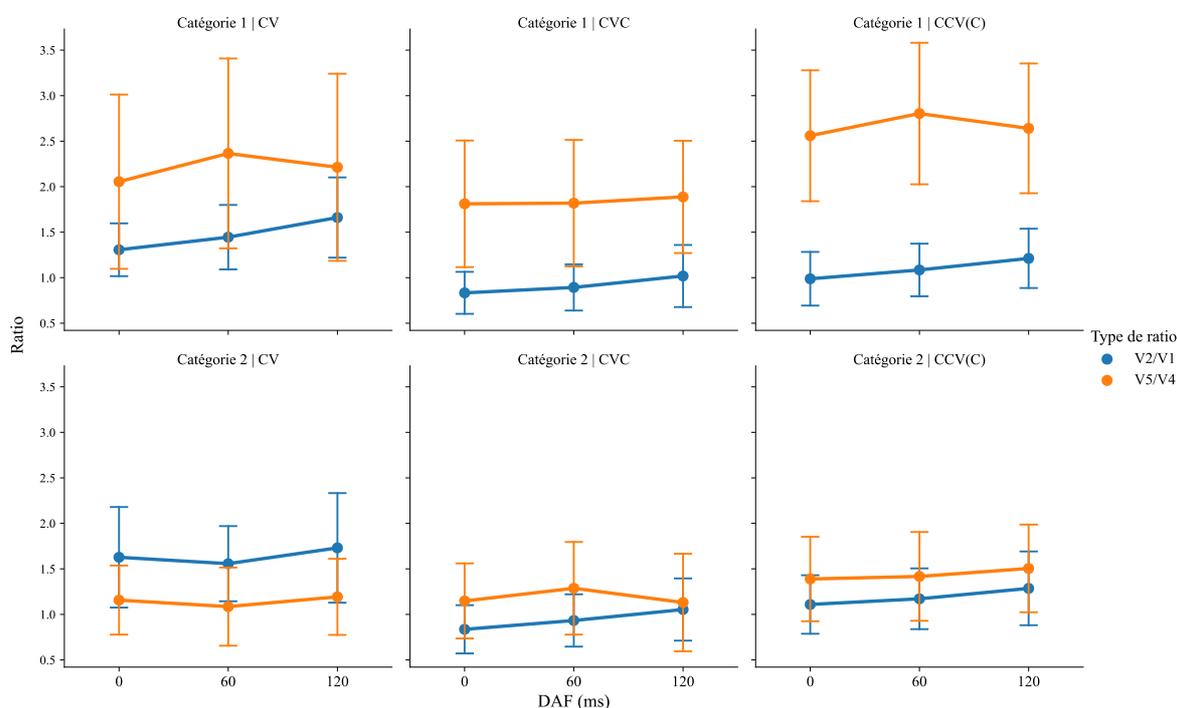
**Résultats de la catégorie d'analyse 2 (phrases finales) :**

Les résultats de la catégorie d'analyse 2 indiquent que la phrase CV dans les essais sans DAF présente un ratio V2/V1 plus élevé que la moyenne de tous les types de phrases. Cependant, la différence entre le ratio V5/V4 de la phrase CV et le ratio V5/V4 moyen pour tous les types de phrases est significativement réduite. En général, le ratio V5/V4 ne diffère pas significativement du ratio V2/V1, à l'exception de la phrase CCV(C) où le ratio V5/V4 est significativement plus élevé que le ratio V2/V1.

En ce qui concerne l'effet du DAF, celui à 60 ms augmente le ratio V2/V1 presque de manière significative pour tous les types de phrases, mais cet effet est affaibli dans la phrase CV. À 120 ms, le DAF augmente de manière significative le ratio V2/V1 pour tous les types de phrases, mais cet effet est affaibli pour le ratio V5/V4.

Figure 17 Moyenne et écart-type du ratio entre la durée de la voyelle accentuée et celle de la voyelle non accentuée dans chaque SA (V2/V1 et V5/V4) pour les locutrices dans la « condition uniquement DAF » dans l'expérience II, respectivement pour chaque catégorie d'analyse (cf. chapitre III 3.1)

Catégorie 1 : phrases non finales ; catégorie 2 : phrases finales



**Conclusion :**

En résumé, dans une certaine mesure, les résultats de l'expérience I (section 2.1) et de l'expérience II (section 2.2) sont cohérents. Le DAF a tendance à allonger davantage les voyelles qui sont accentuées. Cependant, dans l'expérience II, lorsque l'effet de l'accentuation est analysé au

niveau du SA, nous observons un manque d'effet dans les SA à la frontière droite des SIs qui se terminent par un ton montant lorsque le DAF est à un niveau élevé (120 ms). Autrement dit, dans ce SA, sous DAF à 120 ms, la voyelle accentuée n'est pas plus allongée que la voyelle non accentuée.

Nous émettons l'hypothèse que ce manque d'effet soit dû à l'interaction entre l'effet d'allongement final (potentiellement renforcé par le ton montant) sur la dernière syllabe dans les phrases non finales et l'effet d'allongement du DAF. Les analyses visant à explorer sur cette hypothèse seront présentées dans la section suivante (section [2.3](#)).

### 2.3 Effets d'allongement du DAF interagissant avec l'allongement final

Les analyses exposées dans la section [2.2](#) démontrent que la voyelle accentuée n'est pas davantage allongée par le DAF à 120 ms que la voyelle non accentuée dans les phrases non finales. Nous avons émis l'hypothèse que cette absence de différence découle d'une interaction entre l'effet du DAF et l'allongement final à la frontière droite d'un SI.

La Figure [18](#) illustre la durée des voyelles accentuées (V2 dans le SA initial de toutes les phrases ; V5 dans le SA final des phrases non finales ; V5 dans le SA final des phrases finales) dans les essais sans DAF. Il est clairement observable que la voyelle accentuée dans le SA final des phrases non finales (V5 non finale) est considérablement plus longue que les voyelles accentuées dans les autres positions.

Nous posons l'hypothèse que lorsque la voyelle fait déjà l'objet d'un allongement final significatif, un allongement excessif dans le contexte du DAF ne serait plus nécessaire. En d'autres termes, une fois que la durée des syllabes atteint un certain seuil, elle cesserait d'augmenter. En effet, lorsque le DAF est à 60 ms, la durée de la V5 a tendance à augmenter davantage que celle de la V4 dans les phrases non finales (voir Figure [17](#), où le ratio V5/V4 dans les phrases non finales augmente lors que le DAF est à 60 ms). Cependant, lorsque le DAF est à 120 ms, bien que la durée de la V5 augmente par rapport par rapport à celle observée à 60 ms (voir Figure [19](#), qui présente la durée de la V4 et de la V5 à différents niveaux de DAF dans les phrases non finales), la différence de durée entre la V5 et la V4 ne s'accroît pas. Cela explique l'absence d'effet du DAF à 120 ms sur le ratio V5/V4 dans les phrases non finales (voir Figure [17](#)). Le rôle de l'accentuation dans la

modulation de l'effet du DAF pourrait ainsi être neutralisé par l'allongement final à la frontière droite du SI.

Figure 18 Durée des voyelles accentuées dans différentes positions des phrases

V2 : la voyelle accentuée dans le SA initial des phrases ; V5 non finale : la voyelle accentuée dans le SA final des phrases non finales ; V5 finale : la voyelle accentuée dans le SA final des phrases finales

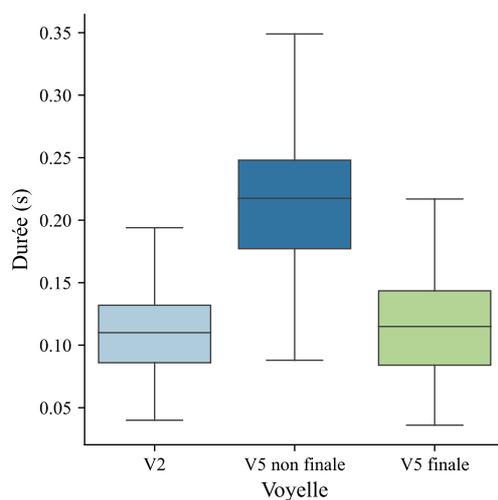
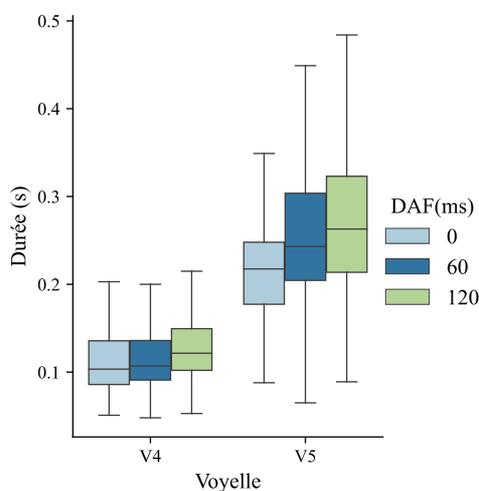


Figure 19 Durée des voyelles V4 et V5 dans les phrases non finales (catégorie d'analyse 1) à divers niveaux DAF (distingués par la couleur)



### 3. Discussion du chapitre

Les analyses effectuées et les résultats présentés dans ce chapitre corroborent notre première hypothèse selon laquelle l'accentuation module les processus de contrôle sensorimoteur. Plus

précisément, lorsque le DAF était appliqué, les locutrices ont montré une tendance à allonger davantage la voyelle accentuée par rapport à la voyelle non accentuée au sein de d'un syntagme accentuel (voir section 2.2). En plus, l'expérience II a permis d'étudier l'effet de l'accentuation dans les énoncés présentant des syllabes de complexité variable. Ainsi, il a été mis en évidence que les différences de durée des noyaux syllabiques engendrées par la structure des syllabes hôtes ne permettaient pas de prédire l'effet du DAF. Cette observation indique que la différence de durée entre les voyelles accentuées et non accentuées ne suffit pas à expliquer la modulation des effets du DAF en fonction de la position structurale des voyelles. En d'autres termes, les syllabes accentuées semblent faire preuve d'une plus grande flexibilité dans le contrôle sensorimoteur. Cette observation permet d'expliquer les résultats de l'étude menée par Chesters et al. (2015), où une variation plus importante de la durée des voyelles a été observé lors d'une tâche de répétition de phrases sous DAF.

Par ailleurs, il est à noter que l'ampleur de l'effet d'allongement du DAF ne peut pas toujours être prédite uniquement en fonction de l'accentuation, car cet effet peut être annulé par l'effet d'allongement final observé à la frontière droite des SIs. En d'autres termes, lorsqu'une syllabe atteint déjà une durée adéquate, l'allongement cesse, même en présence d'un DAF à un niveau élevé.

De plus, les résultats de l'expérience I ont révélé que les effets du DAF ont atteint leur point culminant lorsque le DAF était à 120 ms - 150 ms, indépendamment de la difficulté de production de la phrase. Cette plage de DAF se rapprochait de la durée moyenne des voyelles observée dans les données actuelles (c'est-à-dire 124,433 ms). Des études (ex. Zimmermann et al., 1988) ont également montré que les locuteurs étaient plus perturbés par le DAF lorsque la valeur du DAF était proche de la durée syllabique (c'est-à-dire que le signal retardé est plus susceptible d'être activé pendant le début du mouvement pour le prochain geste).



---

# CHAPITRE V : ANALYSE DU ROLE DU RYTHME DANS LA STABILISATION DU CONTROLE SENSORIMOTEUR PENDANT LA PRODUCTION DE LA PAROLE

## Sommaire du chapitre

1. Effet du DAF sur la coordination des composantes rythmiques dans les phrases de différents niveaux de difficulté de production .....	92
2. Effet du DAF sur la coordination des composantes rythmiques dans les phrases avec différente complexité syllabique .....	94
2.1 Coordination des composantes rythmiques au niveau de la phrase .....	94
2.2 Coordination des composantes rythmiques au niveau du syntagme accentuel .....	96
2.3 Observations complémentaires : variation entre les locutrices .....	98
3. PLV et groupement syllabique .....	101
3.1 Nombre de pics syllabiques et supra-syllabiques .....	101
3.2 Effet du DAF lié au nombre de pics .....	103
4. Discussion du chapitre .....	105
4.1 Établissement de la stabilité de coordination des composantes rythmiques .....	105
4.2 Coordination des composantes rythmiques dans différents syntagmes accentuels .....	107
4.3 Variabilité individuelle .....	108
4.4 Implications des résultats de l'effet du DAF sur l'organisation rythmique .....	109
4.5 Discussion sur la méthode de PLV .....	110

Dans ce chapitre, nous présenterons les résultats de nos analyses visant à explorer notre [deuxième hypothèse](#) (Les perturbations du retour auditif, telles que le DAF, induiraient une réorganisation de la coordination des composantes rythmiques). Nous examinerons les résultats de l'expérience I ainsi que de la « condition uniquement DAF » de l'expérience II (cf. chapitre III description des expériences [2.1](#) et [2.2](#)), et les comparerons si nécessaire.

La section [1](#) de ce chapitre se penchera sur les résultats de l'expérience I, nous permettant de comprendre comment le DAF affecte la coordination entre la composante syllabique et la composante supra-syllabique dans des phrases de différents niveaux de difficulté de production

(par exemple, une phrase normale par rapport à un virelangue, cf. chapitre III [2.1.3](#) Matériaux). Par la suite, dans la section [2](#), nous examinerons les résultats de l'expérience II, en nous concentrant sur l'effet du DAF sur la coordination entre la composante syllabique et la composante supra-syllabique dans des phrases présentant une complexité syllabique variable.

La coordination entre la composante syllabique et la composante supra-syllabique sera quantifiée à l'aide du *Phase Locking Value* (PLV), dont la méthode de calcul a été détaillée dans la section [3.2](#) du chapitre III. Enfin, dans la section [3](#), nous discuterons de nos observations sur le lien entre la PLV et le groupement syllabique tel qu'il ressort de nos résultats.

## **1. Effet du DAF sur la coordination des composantes rythmiques dans les phrases de différents niveaux de difficulté de production**

Dans l'expérience I, nous avons utilisé deux types de phrases comme stimuli. Par exemple, la phrase 1 (/valmɔ̃ vule vwar lə vɔl/), qui est une phrase normale en français avec un patron d'attaque régulier, et la phrase 3 (/mamã e mami mani ma mẽ/), qui est un virelangue en français, présentant une répétition des séquences /mVmV/ et /mVnV/ (cf. chapitre III [2.1.3](#) Matériaux). Ces deux types de phrases introduisent différents niveaux de difficulté dans la production. Comme nous l'avons montré dans la section [2.1.5](#) du chapitre III, les locutrices avaient tendance à commettre plus d'erreurs lors de la production de la phrase 3, et le nombre d'erreurs augmentait avec l'augmentation du niveau du DAF. Par conséquent, dans notre analyse de la PLV, nous nous intéressons non seulement à l'effet du DAF sur la relation temporelle entre les composantes rythmiques, mais également à savoir si cette relation temporelle est modulée par les disfluences causées par la difficulté à produire une phrase (ex. un virelangue). Si les disfluences ont un effet, nous nous attendons à ce que dans la phrase 3, qui a été produite avec plus de disfluences même sans DAF (cf. Figure [10](#)), la coordination entre les composantes rythmiques soit plus forte.

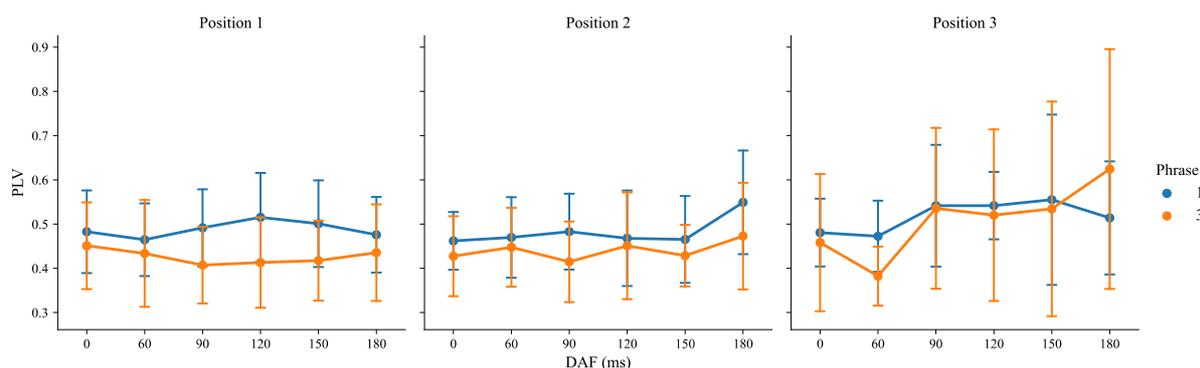
La Figure [20](#) affiche la moyenne et l'écart-type de la PLV à chaque niveau du DAF (0, 60, 90, 120, 150 ou 180 ms) pour la phrase 1 (cf. les courbes en bleu) et la phrase 3 (cf. les courbes en orange). Afin d'estimer l'effet du DAF sur la PLV, un *modèle linéaire mixte* a été conduit. Les prédicteurs étaient le niveau du DAF (référence : 0 ms), la phrase (référence : phrase 1), et l'ordre

de production de la phrase dans un essai (référence : position 1), ainsi que les interactions entre le niveau du DAF et l'ordre de production, et entre la phrase et l'ordre de production. Les données des essais de base n'ont pas été incluses dans le modèle, qui comprenait un total de 910 observations (les résultats clé des modèles cf. le Tableau 6 dans l'Annexe).

### **Résultats :**

La PLV de la phrase 3 est inférieure à celle de la phrase 1 en position 1. Toutefois, cette disparité s'atténue en position 3. En position 1, la présence du DAF n'affecte pas de manière significative la PLV. Néanmoins, le DAF à 180 ms a tendance à augmenter la PLV des répétitions des phrases en positions 2 et 3, tandis que les DAF à 90 ms et 150 ms ont également tendance à augmenter la PLV des répétitions des phrases en position 3. Il est important de noter que ces effets du DAF ne présentent pas de différence significative entre les deux phrases.

*Figure 20 Moyenne et écart-type de la PLV pour les phrases 1 (/valmõ vule vvar læ vl/) et 3 (/mamã e mamì manì ma mẽ/) dans chaque position des essais expérimentaux dans l'expérience I*



### **Conclusion :**

Initialement (c'est-à-dire en position 1 dans la condition sans DAF), la PLV de la phrase plus difficile à produire (la phrase 3, pendant la production de laquelle plus de disfluences ont été produites) est plus faible que celle de la phrase relativement moins difficile à produire (la phrase 1). Comme la phrase 3 a été produite avec plus de disfluences (cf. Figure 10), il est possible que les disfluences n'aient pas induit une PLV plus élevée. Au contraire, la difficulté de la tâche de production peut avoir entraîné une parole moins stable, c'est-à-dire une coordination moins forte entre les composantes rythmiques. Néanmoins, la différence de PLV entre la phrase 1 et la phrase 3 diminue en fin d'un essai expérimental (position 3), ce qui pourrait signifier qu'au fur et à mesure de la production d'un flux de parole, la PLV de la phrase 3 augmente naturellement.

L'effet du DAF interagit avec la position des phrases. Le DAF n'augmente pas immédiatement la PLV des phrases dès le début d'un essai expérimental (en Position 1). En d'autres termes, sous l'effet du DAF, le renforcement de la stabilité de la relation temporelle entre les composantes rythmiques s'établit progressivement. L'effet du DAF ne distingue pas entre la phrase 1 et la phrase 3. Ainsi, l'effet du DAF sur la PLV est indépendant l'effet des disfluences (la difficulté dans la production).

## **2. Effet du DAF sur la coordination des composantes rythmiques dans les phrases avec différente complexité syllabique**

### **2.1 Coordination des composantes rythmiques au niveau de la phrase**

La Figure 21 présente la moyenne et l'écart-type de la PLV à chaque niveau du DAF (0, 60 ou 120 ms) pour chaque type de phrase (CV, CVC, CCV(C), distingués par la couleur). *Un modèle linéaire mixte* a été conduit afin d'estimer l'effet du DAF sur la PLV des phrases. Les prédicteurs du modèle étaient le niveau du DAF (référence : 0 ms), le type de phrase, l'ordre de production de la phrase dans un essai (référence : position 1), ainsi que leurs doubles et triples interactions. L'effet du type de phrase a été codé par une différence par rapport au contraste moyen. Les données des essais de base n'ont pas été incluses dans le modèle, qui comprenait un total de 2682 observations (les résultats clé des modèles cf. le Tableau 7 dans l'Annexe).

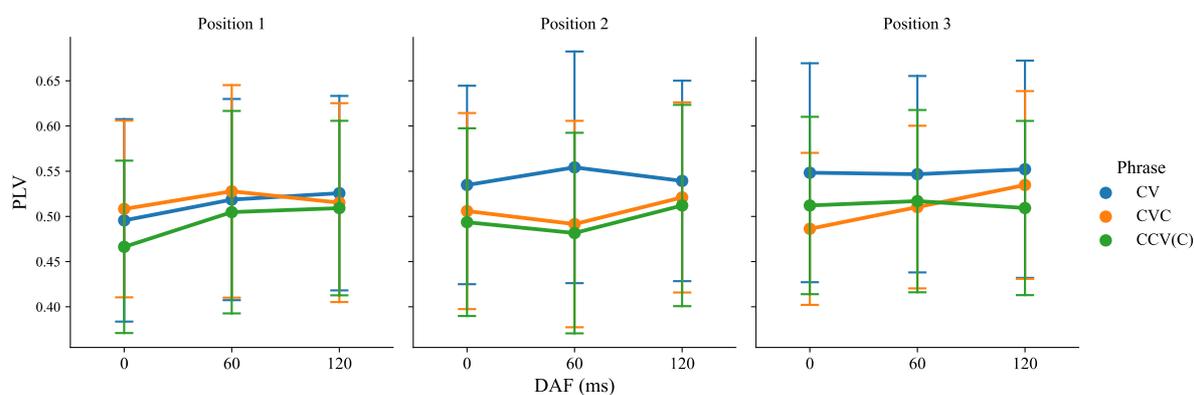
#### **Résultats :**

La PLV de la phrase CCV(C) (cf. Figure 21 les courbes en vert) en position 1 est inférieure à la PLV moyenne des trois types de phrases dans cette position dans la condition sans DAF. En général, la PLV des phrases en positions 2 et 3 est supérieure à celle en position 1 en l'absence du DAF. Pour la phrase CV (cf. Figure 21 les courbes en bleu), la différence de PLV entre les positions 1 et 3 est plus importante.

Par ailleurs, il est à noter que le DAF (60 ou 120 ms) augmente la PLV moyenne des trois types de phrases en position 1. Cependant, son effet a tendance à être atténué en position 2 et position

3, particulièrement observable pour l'effet du DAF à 60 ms en position 2, et pour l'effet du DAF à 120 ms sur la phrase CCV(C) en position 3.

Figure 21 Moyenne et écart-type de la PLV pour chaque type de phrase dans chaque position des essais expérimentaux pour les locutrices dans la « condition uniquement DAF ») dans l'expérience II



### **Conclusion :**

Initialement (c'est-à-dire en position 1 dans la condition sans DAF), la PLV des phrases à structure syllabique complexe (CCV(C)) est plus basse. Dans une certaine mesure, cette observation est cohérente avec les résultats de l'expérience I, où la PLV des phrases plus difficile à produire (phrase 3) est initialement plus basse que celle des phrases relativement faciles à produire (cf. chapitre V section 1). Tout comme dans l'expérience I, la PLV des phrases tend à augmenter naturellement au fur et à mesure de la production du flux de parole sans l'effet du DAF. De plus, cette augmentation est spécifiquement plus prononcée pour la phrase CV. Cela suggère qu'il pourrait être plus facile pour une phrase à structure syllabique simple de renforcer la coordination entre les composantes rythmiques pendant la production.

Cependant, contrairement à ce qui est observé dans l'expérience I, dans l'expérience II, le DAF augmente la PLV des phrases dès le début des essais expérimentaux (position 1). Cette différence entre les deux expériences pourrait être due à la différence de complexité des stimuli utilisés. Les phrases de stimuli de l'expérience II sont plus simples en syntaxe et plus courtes que celles de l'expérience I.

De plus, l'effet du DAF sur les phrases en position 2 et 3 pourrait être atténué par le renforcement naturel de la coordination entre les composantes syllabiques au fur et à mesure de la production du flux de parole. En d'autres termes, une fois qu'une stabilité suffisante est établie, il se peut que le degré de coordination n'ait plus besoin d'être renforcé.

## 2.2 Coordination des composantes rythmiques au niveau du syntagme accentuel

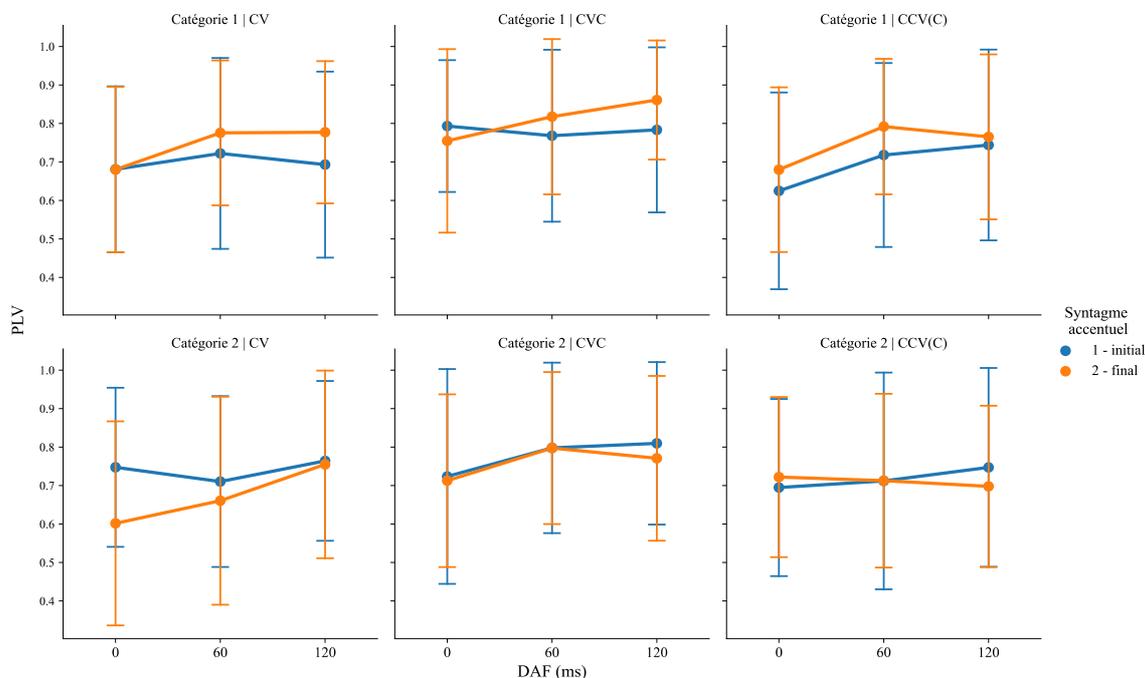
Pour une comparaison plus précise des effets du DAF sur la PLV avec ceux du DAF sur le ratio de la durée de la voyelle accentuée et de la voyelle non accentuée dans chaque SA (cf. chapitre IV [2.2](#)), nous avons calculé une PLV dans chaque SA de chaque phrase. Plus précisément, nous avons calculé la PLV moyenne en prenant en compte les valeurs de phase observées dans la portion allant du début de l'énoncé à la fin de la deuxième voyelle pour le SA initial et en prenant en compte les valeurs de phase observées dans la portion allant du début de la quatrième voyelle à la fin de l'énoncé pour le SA final. Les PLVs résultant des deux SAs ont été regroupées en deux catégories d'analyse (cf. chapitre III [3.1](#), comprenant soit des phrases non finales avec des tons montants finaux, soit des phrases finales avec des tons descendants finaux).

Il convient de noter que cette section ne constitue qu'une analyse exploratoire. L'inconvénient de cette analyse est évident : le calcul de la PLV repose sur la méthode d'analyse des modulations d'amplitude (cf. chapitre III [3.2](#)), qui reflète les fluctuations d'énergie dans le signal acoustique. Les prééminences identifiées par la méthode de filtrage passe-bande peuvent ne pas correspondre aux accents définis sur le plan phonologique (cf. Figure [12](#)). Le calcul de la PLV dans chaque SA présuppose la possibilité d'identifier au moins une oscillation complète de la composante supra-syllabique dans une fenêtre d'analyse prédéfinie (c'est-à-dire, un SA). Cependant, malgré cet inconvénient, dans une certaine mesure, la PLV calculée dans chaque SA des phrases permet de rendre compte de la stabilité de la relation temporelle des composantes identifiées dans chaque SA.

La Figure [22](#) présente la moyenne et l'écart-type de la PLV à chaque niveau du DAF (0, 60 ou 120 ms) pour le SA initial et le SA final (distingués par la couleur) dans chaque type de phrase (CV, CVC, CCV(C), correspondant à chaque colonne). *Deux modèles linéaires mixtes* ont été conduits séparément pour la catégorie d'analyse 1 et la catégorie d'analyse 2. Les prédicteurs utilisés étaient le niveau du DAF (référence : 0 ms), le SA (référence : SA initial), le type de phrase, ainsi que leurs doubles et triples interactions. L'effet du type de phrase a été codé en utilisant une déviation du contraste moyen. Les données des essais de base n'ont pas été incluses dans les modèles. Le modèle pour la catégorie d'analyse 1 comprenait un total de 1800 observations, tandis que le modèle pour la catégorie d'analyse 2 comprenait un total de 1247 observations (les résultats clé des modèles cf. le Tableau [8](#) dans l'Annexe).

Figure 22 Moyenne et écart-type de la PLV dans chaque SA pour les locutrices dans la « condition uniquement DAF » dans l'expérience II, respectivement pour chaque catégorie d'analyse

Catégorie 1 : phrases non finales ; Catégorie 2 : phrases finales



### Résultats de la catégorie d'analyse 1 (phrases non finales) :

Dans la catégorie d'analyse 1, la PLV du SA initial est plus faible pour la phrase CCV(C) produites dans la condition sans DAF, comparée à la PLV moyenne pour les trois types de phrases. Les deux niveaux du DAF n'entraînent pas une augmentation significative de la PLV moyenne dans le SA initial des trois types de phrase. En revanche, ils provoquent une augmentation significativement plus importante de la PLV dans le SA initial de la phrase CCV(C). Dans l'ensemble, **l'effet du DAF sur la PLV est plus important dans le SA final** pour tous les types de phrases, à l'exception de l'effet du DAF à 120 ms sur la PLV dans le SA final de la phrase CCV(C).

### Résultats de la catégorie d'analyse 2 (phrases finales) :

Dans la catégorie d'analyse 2, la PLV de la phrase CV est inférieure dans le SA final par rapport au SA initial en l'absence du DAF. Cette différence entre la PLV du SA final et du SA initial est plus faible dans la phrase CCV(C). Le DAF a tendance à augmenter la PLV moyenne du SA initial pour les trois types de phrases, ainsi que la PLV moyenne du SA final pour les trois types de phrases. Cependant, de manière générale, **l'effet du DAF n'est pas significatif**. Seul le DAF à 120 ms augmente de manière significative la PLV du SA final dans la phrase CV.

### **Conclusion :**

En général, les résultats de l'analyse de la PLV au niveau du SA sont cohérents avec ceux de l'analyse de la PLV au niveau de la phrase (cf. chapitre V [2.1](#)). Tout d'abord, la PLV dans le SA initial de la phrase à la structure syllabique complexe (CCV(C)) peut être initialement plus faible (c'est-à-dire dans les phrases non finales - catégorie d'analyse 1). Cette faible PLV dans le SA initial de la phrase CV, pourrait contribuer à une augmentation précoce de la PLV dans ce SA dès le début d'un essai. Cette observation confirme la nécessité d'un certain degré de coordination entre les composantes rythmiques pour assurer la stabilité de la production de la parole dans la condition du DAF. En revanche, pour les autres types de phrases, l'augmentation de la PLV se produit principalement dans le SA final des phrases non finales. Cependant, la PLV plus élevée dans le SA final que dans le SA initial pourrait également être liée à un effet de la position structurale du SA. Il est possible que le renforcement de la coordination entre les composantes rythmiques soit plus marqué dans les SA situés à la frontière droite d'un SI, ce qui correspond aux résultats des analyses sur l'effet du DAF sur la différence de durée entre les voyelles accentuées et non accentuées.

D'un autre côté, l'effet du DAF sur la PLV est généralement moins prononcé dans les phrases finales (catégorie d'analyse 2), ce qui corrobore l'observation dans l'analyse de la PLV au niveau de la phrase, indiquant que l'effet du DAF sur la PLV dans les phrases à la fin des essais expérimentaux est atténué par le renforcement naturel de la PLV au fur et à mesure de la production d'un flux de parole. Cependant, une forte augmentation de la PLV est observée pour le SA final de la phrase CV en position finale des essais sous l'effet du DAF. Cette augmentation significative pourrait être attribuable au fait que la PLV dans ce SA est initialement très basse. Ainsi, un renforcement substantiel de la coordination est nécessaire pour garantir la stabilité de la production de la parole sous l'effet du DAF. Cependant, l'origine de cette PLV initialement basse dans ce SA demeure inconnue.

### **2.3 Observations complémentaires : variation entre les locutrices**

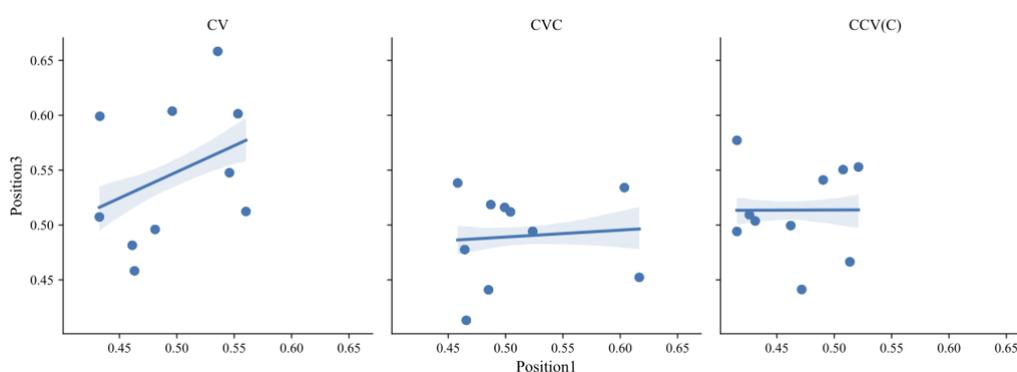
Des études, telles que celle menée par Cichocki (2015), ont montré une grande variation du rythme de la parole entre les locuteurs. Dans les figures présentées dans les sections [2.1](#) et [2.2](#) (les

Figures 21 et 22), nous avons également constaté une grande variation de la coordination entre les composantes rythmiques parmi les participantes. Afin de mieux visualiser cette variation, nous avons créé deux nuages de points (les Figure 23 et 24) dans lesquels chaque point représente la PLV moyenne d'un type de phrase pour une locutrice donnée dans la condition correspondante.

La Figure 23 représente la relation entre la PLV moyenne de chaque locutrice pour les phrases en position 1 (axe X) et en position 3 (axe Y) dans la condition sans DAF (les données des essais de base ont été exclues). Nous constatons que la variation est présente en position 1 et en position 3 pour les trois types de phrase. La variation entre les locutrices tend à être plus importante en position 3 qu'en position 1, en particulier pour la phrase CV, mais à l'exception de la phrase CVC. Parmi les trois types de phrases, la variation entre les locutrices est plus faible pour la phrase CCV(C). De manière générale, la PLV en position 1 et en position 3 ne présente presque pas de relation linéaire. Autrement dit, une locutrice dont la PLV moyenne est moins élevée en position 1 n'a pas nécessairement une PLV moyenne moins élevée en position 3 par rapport aux autres locutrices.

Figure 23 PLV moyenne de chaque type de phrase des répétitions en position 1 (axe X) et en position 3 (axe Y) pour les locutrices dans la « condition uniquement DAF » dans l'expérience II.

Chaque point représente la PLV moyenne d'une locutrice dans la condition correspondante.

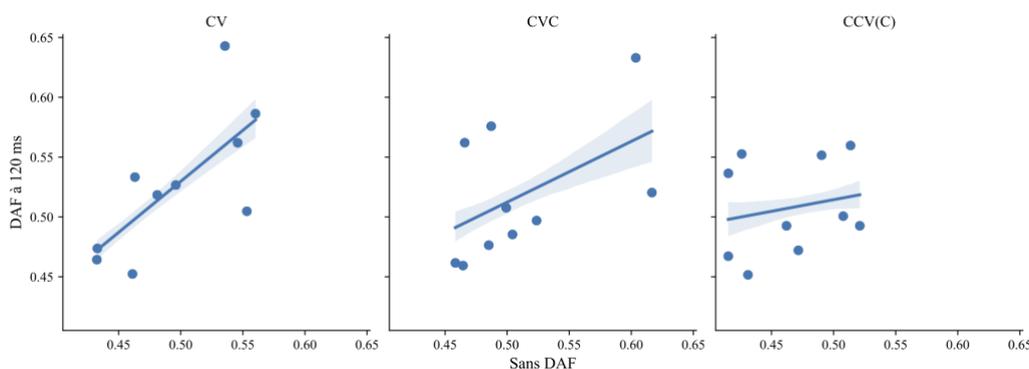


La Figure 24 lie la PLV moyenne de chaque locutrice obtenue dans la condition sans DAF (axe X) à celle obtenue sous DAF de 120 ms (axe Y) dans les répétitions de phrases produites en position 1 des essais (les données des essais de base ont été exclues). Nous observons une variation importante entre les locutrices, en particulier pour la phrase CV où la variation est plus grande sous DAF à 120 ms que dans la condition sans DAF. En revanche, pour la phrase CCV(C), la variation reste la plus petite parmi les trois types de phrases et ne change pas beaucoup sous DAF. La PLV obtenue dans la condition sans DAF et celle obtenue sous DAF à 120 ms ont tendance à avoir une relation linéaire, ce qui signifie que les locutrices ayant une PLV moyenne moins élevée dans la

condition sans DAF ont également tendance à avoir une PLV moyenne moins élevée sous DAF à 120 ms, par rapport aux autres locutrices.

Figure 24 PLV moyenne de chaque type de phrase produite en position 1 pour les locutrices dans la « condition uniquement DAF » dans essais sans DAF (axe X) et dans les essais avec DAF à 120 ms (axe Y) dans l'expérience II.

Chaque point représente la PLV moyenne d'une locutrice dans la condition correspondante.



### **Conclusion :**

Il est essentiel de noter que le degré de coordination entre les composantes rythmiques de la parole peut varier considérablement d'une locutrice à l'autre. Plusieurs observations sont pertinentes en relation avec cette variation.

Tout d'abord, il est à rappeler que les résultats de l'analyse présentée dans la section 2.1 montrent que la *Phase Locking Value* (PLV) d'une phrase tend à augmenter au fur et à mesure de la production. Cependant, cette augmentation de la coordination entre les composantes rythmiques (PLV), ne s'applique pas uniformément à toutes les locutrices (cf. Figure 23).

De plus, dans une certaine mesure, la structure syllabique de la phrase peut influencer la variation de la PLV entre les locutrices. En général, cette variation est moins marquée dans toutes les conditions pour les phrases à structure complexe (CCV(C)).

Enfin, la relation plus ou moins linéaire entre la PLV obtenue dans la condition sans DAF et celle obtenue dans la condition avec DAF à 120 ms (cf. Figure 24) indique que l'augmentation de la coordination entre les composantes rythmiques (mesurée par la PLV) est largement adoptée par la plupart des participantes, à l'exception des phrases à structure syllabique complexe (CCV(C)), pour lesquelles cette relation linéaire est moins évidente.

### 3. PLV et groupement syllabique

La PLV est une mesure qui évalue la stabilité de la relation temporelle entre deux signaux oscillatoires, dans notre cas, l'oscillation syllabique et l'oscillation supra-syllabique (cf. chapitre III [3.2](#)). Bien que nous ayons montré que le degré de coordination entre la composante syllabique et la composante supra-syllabique est lié au groupement syllabique (cf. chapitre I [3.2.2](#)), il est essentiel d'analyser cette relation entre la PLV et le groupement syllabique au cas par cas dans le contexte du DAF. En fait, nous avons remarqué que certaines participantes avaient tendance à produire une parole saccadée sous l'effet du DAF. Cela signifie qu'elles avaient tendance à attribuer la même proéminence à chaque syllabe pour assurer la stabilité de leur production. Dans le cas d'une parole saccadée, nous postulons que la composante syllabique et la composante supra-syllabique extraites par la méthode du filtrage passe-bande seraient très similaires, ce qui entraînerait également une PLV élevée.

#### 3.1 Nombre de pics syllabiques et supra-syllabiques

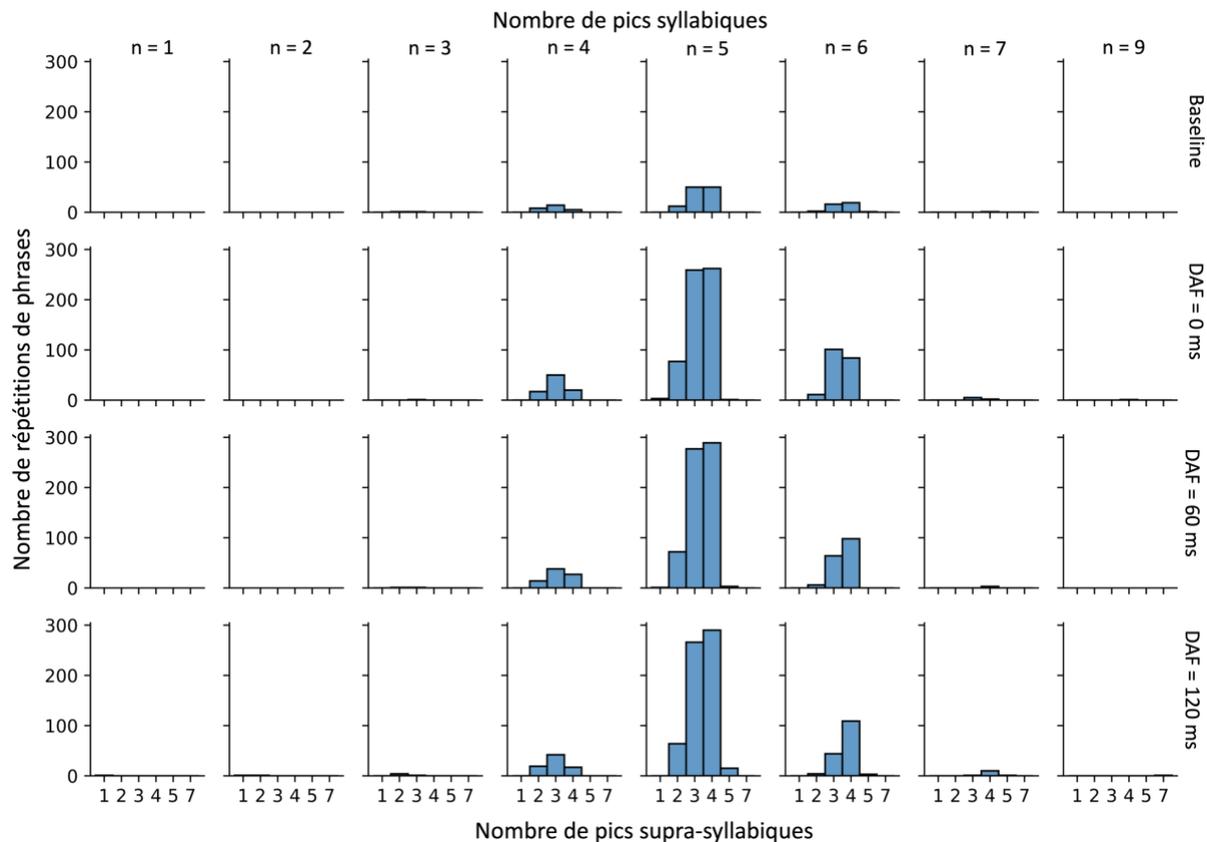
Dans le but de vérifier cette hypothèse, dans la Figure [25](#), nous présentons des histogrammes illustrant les distributions du nombre de répétitions de phrases<sup>9</sup> en fonction du nombre de pics supra-syllabiques. Ces données sont calculées séparément en tenant compte du nombre de pics syllabiques présents dans chaque énoncé (correspondant aux colonnes) et des différentes conditions du DAF (correspondant aux lignes). De cette manière, nous pouvons visualiser la distribution du nombre de répétitions de phrases pour chaque combinaison de ces trois variables, en utilisant les données des participantes dans la « condition uniquement DAF » de l'expérience II (cf. chapitre III [2.2](#)). Par exemple, nous pouvons observer que dans la plupart des répétitions de phrases, cinq pics syllabiques (colonne 5 dans la Figure [25](#)) ont été identifiés, ce qui était attendu puisque les phrases de stimuli de l'expérience II sont phonologiquement composées de cinq syllabes. De plus, le nombre de répétitions de phrases contenant quatre pics syllabiques (colonne 4) ou six pics syllabiques (colonne 6) est également relativement élevé. Le cas de quatre pics syllabiques pourrait correspondre à une situation où deux syllabes consécutives étaient trop

---

<sup>9</sup> Pour rappel, le nombre total de répétitions de phrases réalisées par l'ensemble des participantes s'élève à 2880 dans la « condition uniquement DAF » de l'expérience II (cf. chapitre III [2.2.4](#)).

dépendantes pour être différenciées. En revanche, le cas de six pics supra-syllabiques pourrait refléter une situation où une syllabe particulière de la phrase avait été considérablement étendue, générant ainsi deux pics pendant la production d'une seule syllabe.

Figure 25 Distribution du nombre de répétitions de phrases contenant différents nombres de pics syllabiques (sur les colonnes) et de pics supra-syllabiques (sur l'axe des X) identifiés grâce à notre méthode d'analyse des modulations d'amplitude (cf. chapitre III 3.2), respectivement pour chaque condition du DAF (en colonnes), concernant les locutrices dans la « condition uniquement DAF » de l'expérience II

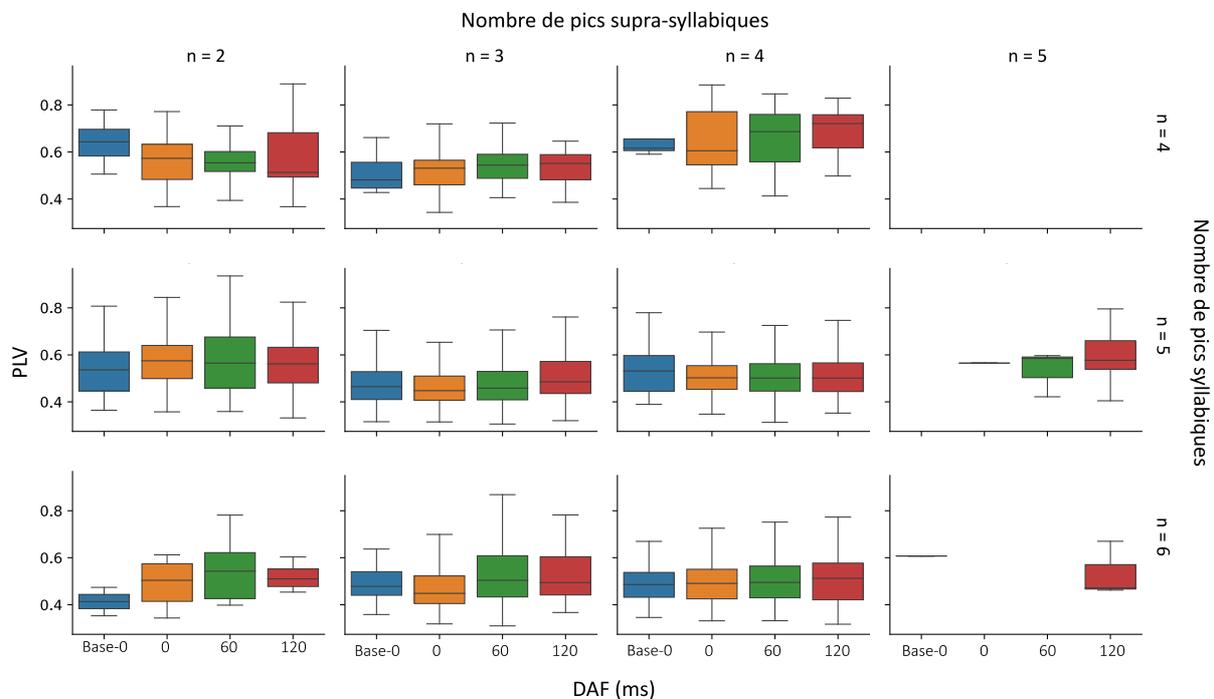


Dans la plupart des répétitions de phrases contenant cinq pics syllabiques, trois ou quatre pics supra-syllabiques ont été identifiés. Un certain nombre de répétitions de phrases ont révélé la présence de deux pics supra-syllabiques. Il est à noter que dans la condition avec un niveau de DAF élevé (120 ms), le nombre de répétitions de phrases contenant cinq pics supra-syllabiques a légèrement augmenté par rapport à la condition sans DAF ou à la condition avec un niveau de DAF moins élevé (60 ms). En d'autres termes, le DAF a favorisé l'émergence d'une parole saccadée. De manière similaire, dans la plupart des répétitions de phrases contenant quatre ou six pics syllabiques, trois ou quatre pics supra-syllabiques ont été observés.

### 3.2 Effet du DAF lié au nombre de pics

La Figure 26 présente la PLV de phrase (cf. chapitre V 2.1) à chaque niveau de DAF (essais de base, 0, 60, 120 ms) pour les répétitions de phrases contenant différents nombres de pics syllabiques (correspondant aux lignes) et différents nombres de pics supra-syllabiques (correspondant aux colonnes). Dans cette figure, seuls le nombre de pics syllabiques (4, 5, 6) et de pics supra-syllabiques (2, 3, 4, 5) le plus fréquemment observés (cf. Figure 25) ont été inclus.

Figure 26 PLV de phrase à chaque niveau du DAF (axe X) dans les répétitions de phrases où différents nombres de pics syllabiques (lignes) et supra-syllabiques (colonnes) ont été trouvés par notre méthode d'analyse des modulations d'amplitude (cf. chapitre III 3.2), pour les locutrices dans la « condition uniquement DAF » de l'expérience II



Avant de considérer l'effet du DAF sur la PLV, il est à noter que dans les répétitions de phrases où le nombre de pics syllabiques est similaire ou identique au nombre de pics supra-syllabiques (par exemple, cinq pics syllabiques vs. quatre pics supra-syllabiques ou quatre pics syllabiques vs. quatre pics supra-syllabiques), la PLV est généralement plus élevée que dans les répétitions de phrases contenant moins de pics supra-syllabiques. Cela peut être vérifié en effectuant une comparaison des « boîtes » bleues de la deuxième rangée dans la Figure 26, lesquels représentent les répétitions de phrases contenant cinq pics syllabiques. Cette observation s'explique par le fait que lorsque le nombre de pics syllabiques et de pics supra-syllabiques est similaire ou identique, les oscillations syllabiques et supra-syllabiques sont susceptibles d'être similaires, ce qui garantit naturellement une

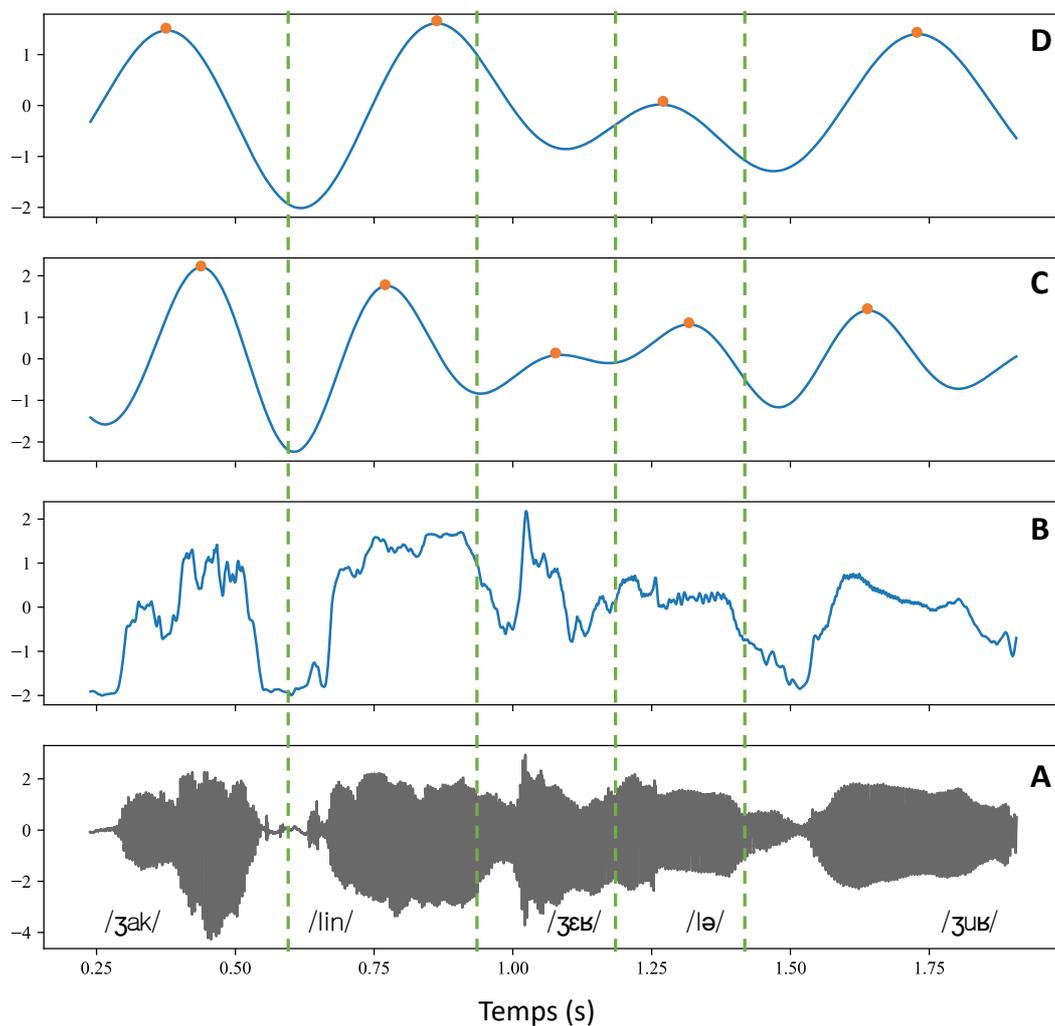
relation temporelle plus stable. À titre d'exemple, la Figure 27 présente les signaux syllabique (panneau C) et supra-syllabique (panneau D) de la phrase /zaklin zεβ lə zυβ/ dans le cas où cinq pics syllabiques et quatre pics supra-syllabiques ont été identifiés. Nous pouvons constater que la forme des signaux syllAM et accAM est similaire.

Figure 27 Modulations d'amplitude pour la phrase « Jacqueline gère le jour » (/zaklin zεβ lə zυβ/)

Exemple où cinq pics syllabiques (panneau C) et quatre pics supra-syllabiques (panneau D) ont été identifié.

A : le signal acoustique ; B : la variation de l'intensité ; C : l'oscillation syllabique ; D : l'oscillation supra-syllabique

Les valeurs Y dans chaque panneau correspondent aux z- scores. Les lignes verticales en vert marquent les frontières des syllabes.



Par ailleurs, l'effet du DAF est moins marqué dans les répétitions de phrases où le nombre de pics syllabiques est similaire ou identique au nombre de pics supra-syllabiques (par exemple, cinq pics syllabiques vs. quatre pics supra-syllabiques ou quatre pics syllabiques vs. quatre pics supra-syllabiques). Cette observation entraîne deux conclusions importantes.

Tout d'abord, l'effet du DAF sur la PLV de phrase, tel que présenté dans la section [2.1](#), est principalement attribuable aux variations de la PLV dans les répétitions de phrases où un nombre plus restreint de pics supra-syllabiques a été identifié (par exemple, dans les répétitions où cinq pics syllabiques coïncident avec deux ou trois pics supra-syllabiques). Dans cette perspective, il semble que sous l'effet du DAF, au moins certaines des participantes aient renforcé le groupement syllabique dans certaines répétitions de phrases.

Ensuite, il convient de noter qu'au sein du contexte du DAF, il n'est pas adéquat de lier simplement la PLV, qui mesure la stabilité de la relation temporelle entre deux signaux, au degré de groupement syllabique. Néanmoins, la PLV demeure une mesure de la rythmicité de la parole. En d'autres termes, une parole fortement rythmique peut résulter d'un renforcement du groupement syllabique, ou de la grande similitude entre les syllabes, bien que cette dernière situation ne corresponde pas nécessairement à une parole naturelle dans le contexte de la vie quotidienne.

## 4. Discussion du chapitre

Les résultats des analyses présentés dans ce chapitre confirment notre [deuxième hypothèse](#), qui stipule que le DAF induirait une réorganisation de la structure rythmique. En effet, la présence du DAF nécessite une coordination plus étroite entre la composante syllabique et la composante supra-syllabique pour garantir la stabilité de la production de la parole. Cependant, il convient de noter que cette coordination est soumise à l'influence de plusieurs facteurs, et une grande variabilité individuelle a été observée. Nous aborderons en premier lieu ces deux aspects. De plus, nous discuterons des implications de cette réorganisation rythmique sous l'effet du DAF. Enfin, nous examinerons la méthode PLV que nous avons utilisée pour quantifier le degré de coordination entre les composantes rythmiques.

### 4.1 Établissement de la stabilité de coordination des composantes rythmiques

Les résultats de l'expérience I (cf. section [1](#)) ont révélé que le degré de coordination entre les deux composantes rythmiques était associé au niveau de la difficulté des tâches de production de

la parole. En l'absence du DAF, la PLV était plus élevée pour la phrase 1, qui est une phrase assertive normale, que pour la phrase 3, qui est un virelangue plus difficile à produire. Ces résultats suggèrent que le renforcement de la coordination entre les composantes rythmiques ne peut pas expliquer le degré de disflueance résultant du DAF. Si cela était le cas, les niveaux de coordination observés devraient être plus élevés dans la phrase 3, qui présente davantage d'erreurs de parole (disfluences), que dans la phrase 1.

De manière similaire, les résultats de l'expérience II (cf. [2.1](#)), ont mis en évidence que la phrase CCV(C), caractérisée par une structure syllabique plus complexe, a présenté une PLV plus faible en comparaison à la PLV moyenne observée pour les trois types de phrases (CV, CVC, CCV(C)) dans la condition sans DAF. Cela suggère que le degré de coordination entre les composantes rythmiques pourrait être lié à la structure syllabique.

Les résultats des deux expériences ont mis en évidence un renforcement naturel de la coordination entre les composantes rythmiques au fur et à mesure de la production d'un flux de parole. Dans l'expérience I, la différence de PLV entre les deux phrases diminuait à mesure que la position dans l'essai expérimental augmente. Quant à l'expérience II, la PLV moyenne des trois phrases dans les positions 2 et 3 au sein d'un essai était supérieure à celle de la position 1 (au début de l'essai).

Cependant, il est important de noter que le moment où l'effet du DAF sur la PLV a commencé à se manifester différait entre les deux expériences. Dans l'expérience I, l'effet du DAF sur la PLV s'est établi progressivement au cours d'un essai expérimental. En revanche, dans l'expérience II, l'effet du DAF sur la PLV s'est manifesté dès la première phrase d'un essai. Cette observation peut s'expliquer par l'utilisation de phrases plus courtes avec une structure syntaxique plus régulière dans l'expérience II, ce qui favorise la mise en place rapide d'une coordination stable entre les composantes rythmiques sous l'influence du DAF. En revanche, les phrases de stimuli sont plus complexes dans l'expérience I, nécessitant ainsi davantage de temps pour établir une coordination suffisamment stable sous l'effet du DAF.

De plus, lorsque le degré de coordination entre les composantes rythmiques est déjà élevée pour assurer une production stable dans la condition du DAF, un renforcement supplémentaire de cette coordination devient moins nécessaire. Par exemple, dans l'expérience II, les phrases produites en positions 2 et 3, qui présentent déjà une PLV plus élevée en l'absence de DAF, montrent un effet moindre du DAF sur la PLV de ces phrases. Cependant, lorsque le niveau de cette coordination

n'est pas suffisant, les locutrices peuvent déployer des efforts supplémentaires pour améliorer la coordination. Par exemple, dans la condition sans DAF, la PLV du SA initial dans la phrase CCV(C) est inférieure à la PLV moyenne du SA initial des trois types de phrases, ce qui expliquerait pourquoi l'effet du DAF se traduit par une augmentation plus marquée de la PLV du SA initial dans la phrase CCV(C) par rapport aux autres phrases.

## 4.2 Coordination des composantes rythmiques dans différents syntagmes accentuels

L'analyse de la coordination entre les composantes rythmiques au sein de chaque SA dans l'expérience II semble indiquer que le DAF renforce davantage la coordination des composantes rythmiques dans le SA final que dans le SA initial, pour les phrases produites non finales au sein d'un essai. Autrement dit, l'effet du DAF sur la coordination entre les composantes rythmiques peut varier en fonction de la position du SA. Il semble que la coordination entre les composantes rythmiques à un niveau prosodique plus élevé (à la frontière droite d'un SI) soit plus importante en présence du DAF.

Cependant, comme évoqué dans la section [2.2](#), cette analyse ne peut être qu'exploratoire. En effet, le calcul de la PLV dans chaque SA suppose l'existence d'au moins une oscillation complète de la composante supra-syllabique au sein de chaque SA prédéfini (c'est-à-dire le SA initial ou le SA final, cf. chapitre III [3.1](#)), ce qui n'est pas nécessairement le cas dans toutes les situations. Les oscillations de la composante supra-syllabique émergent de la production des syllabes les plus proéminentes (cf. chapitre III [3.2.1](#)). Cependant, leur détection est basée sur la méthode d'analyse des modulations d'amplitude, qui repose exclusivement sur la variation de l'énergie dans le signal acoustique. Par conséquent, nous ne pouvons pas postuler une correspondance biunivoque entre les oscillations observées dans le signal supra-syllabique et les accents (cf. Figure [12](#)), en particulier dans le cas du français, où l'intensité joue un rôle moins prépondérant dans l'accentuation (Delattre, 1966b). En d'autres termes, bien que l'accentuation favorise l'émergence des syllabes physiquement saillantes (en particulier dans les langues où l'intensité prédit efficacement les emplacements des accents), conduisant au groupement des syllabes, il est important de noter que le groupement physique des syllabes ne dépend pas uniquement des accents.

### 4.3 Variabilité individuelle

Dans la condition sans DAF, une grande variabilité individuelle a été observée en ce qui concerne le degré de coordination entre les composantes rythmiques, ainsi que le renforcement de cette coordination au fil de la production d'un flux de parole (cf. [2.3](#)). Cette variabilité significative pourrait être en partie liée à la variation du débit de parole entre les participantes. En effet, notre étude récente (Lancia et al., 2023) a révélé qu'un débit de parole plus élevé favorisait l'établissement d'une coordination plus stable entre les composantes rythmiques.

De plus, la variabilité individuelle du degré de coordination entre les composantes rythmiques semble être plus marquée pour les phrases présentant une structure syllabique simple (CV) par rapport à celles ayant une structure syllabique complexe (CCV(C)). Les origines de cette différence n'ont pas été explorées dans le cadre de cette thèse. Cependant, nous émettons l'hypothèse qu'une relation temporelle plus flexible entre les composantes rythmiques peut être à l'origine de cette variabilité, étant donné que le renforcement de la coordination entre les composantes rythmiques est plus prononcé pour la phrase CV, que ce soit au fil de la production du flux de parole (cf. [Figure 23](#)) ou sous l'effet du DAF (cf. [Figure 24](#)).

De plus, bien que dans l'ensemble, l'effet du DAF renforce la coordination entre les composantes rythmiques, nous avons observé différentes stratégies adoptées par les locutrices pour atteindre ce renforcement. D'une part, certaines locutrices montraient la tendance à produire une parole saccadée en présence du DAF, ce qui se traduisait par une attribution de proéminence similaire à chaque syllabe, entraînant ainsi une similitude accrue entre la composante syllabique et la composante supra-syllabique. Ce phénomène conduisait à une coordination élevée entre ces deux composantes. Par conséquent, le renforcement de cette coordination n'était plus nécessaire avec l'augmentation du DAF (cf. [Figure 26](#)).

D'autre part, certaines locutrices adoptaient la stratégie de renforcer le degré du groupement syllabique ou la régularité du groupement syllabique en présence du DAF, ce qui se traduisait par une augmentation de la PLV dans les essais avec DAF par rapport aux essais sans DAF (cf. [Figure 26](#)). Ces locutrices avaient tendance à grouper leurs syllabes en ensembles de deux ou trois.

#### 4.4 Implications des résultats de l'effet du DAF sur l'organisation rythmique

La réorganisation des patrons rythmiques de la parole sous DAF met en évidence que le rythme de la parole est à la fois spécifique à la langue ou au type rythmique, et dépendant du contexte de l'énonciation. Autrement dit, le rythme de la parole résulte non seulement de la réalisation de propriétés phonologiques, phonétiques et structurelles propres à la langue (cf. chapitre I section 2), contraintes par des facteurs (neuro-)biologiques universels (cf. chapitre I 1.3) et d'autres facteurs non linguistiques (ex. le style d'élocution, cf. Leong et Goswami, 2015), mais également du traitement des informations sensorimotrices pendant la production de la parole. En effet, les locuteurs évaluent constamment le couplage sensorimoteur (cf. chapitre II section 1) et adaptent avec souplesse leur comportement aux perturbations (ex. celles qui modifient le retour auditif), comme le montrent les résultats de la présente étude. Le rythme, en tant que mécanisme coordinatif, joue un rôle essentiel dans la stabilisation des processus de contrôle sensorimoteur au cours de la production de la parole.

Par ailleurs, il a été montré que la modification du retour auditif peut considérablement améliorer la fluidité de la parole chez les personnes souffrant de bégaiement (cf. Bradshaw et al., 2021 pour une revue de la littérature). Les résultats de notre étude peuvent apporter des éclaircissements sur les effets de la modification du retour auditif sur la fluidité de la parole chez les personnes atteintes de bégaiement. Étant donné que le bégaiement affecte principalement les aspects rythmiques de la parole, nous pouvons supposer que la modification du retour auditif favorise la réorganisation des patrons rythmiques de la parole de manière à améliorer la fluidité temporelle. Des recherches futures approfondies sur la manière dont cette réorganisation est réalisée et sur ses caractéristiques pourraient être bénéfiques pour la caractérisation du comportement des personnes atteintes de bégaiement ainsi que pour la conception de nouvelles approches thérapeutiques.

Les résultats de ce chapitre suggèrent également que la manipulation du système sensorimoteur pourrait potentiellement favoriser l'acquisition d'un patron rythmique spécifique. En effet, les enfants de certaines langues maîtrisent le rythme de la parole de type adulte plus tard que les enfants d'autres langues en raison de la complexité de la structure rythmique de certaines langues par rapport à d'autres (Grabe et al., 1999). Par exemple, l'anglais a un placement d'accent moins prévisible, des paramètres phonétiques plus complexes pour les accents et une durée de syllabe plus variable que le français (Grabe et al., 1999). Des recherches ont montré que la structuration

du rythme de la parole à la manière des adultes est déjà en place chez les enfants français de 2 ans (Hallé et al., 1991), tandis qu'elle peut être encore décalée chez les enfants anglais de 6 ans (Payne et al., 2012). Le fait que le rythme de la parole soit une propriété émergente de la production de la parole suggère la possibilité de faciliter l'acquisition du langage en contrôlant les conditions de production de la parole. À titre d'exemple, il est légitime de postuler que les enfants éprouvant des difficultés à acquérir une langue avec des patrons rythmiques complexes, ou plus généralement, les enfants présentant des disfluences dans leur parole, pourraient tirer profit d'une exposition à des altérations du retour auditif ou à la production de phrases favorisant l'émergence des patrons rythmiques souhaités. Cette approche peut même être généralisée pour stimuler l'apprentissage d'une seconde langue chez les adultes.

Il est pertinent de noter que, bien que des études développementales (ex. Grabe et al., 1999), aient suggéré que l'apprentissage d'une langue dotée de caractéristiques rythmiques complexes, souvent marquée par une forte coordination entre la composante syllabique et la composante supra-syllabique (ex. l'anglais, cf. Lancia et al., 2019), puisse présenter des défis, la présente étude révèle certains avantages potentiels liés à l'adoption d'une organisation rythmique plus complexe. Cette organisation, grâce à la coordination étroite entre les composantes rythmiques, permet un contrôle plus stable de la production de la parole.

#### 4.5 Discussion sur la méthode de PLV

Nous avons employé la méthode de la *Phase Locking Value* (PLV) pour quantifier le degré de coordination entre la composante syllabique et la composante supra-syllabique. L'un des avantages de cette méthode, qui repose sur la technique de modulation d'amplitude, est son automatisation de l'ensemble des processus d'analyse (cf. chapitre III 3.2). En d'autres termes, le calcul de la PLV n'exige pas nécessairement des segmentations manuelles des unités linguistiques.<sup>10</sup> En ajustant la fréquence des filtres, il est possible d'isoler automatiquement les composantes rythmiques à

---

<sup>10</sup> Dans le cadre de cette thèse, les segmentations des voyelles ont été utilisées afin d'améliorer la normalisation des signaux avant de calculer la PLV dans cette étude (cf. chapitre III 3.2.4). Cependant, les segmentations ont été réalisées par une méthode automatique qui se base sur le *Dynamic Time Warping* (Sakoe et Chiba, 1978). Pour l'expérience I, les segmentations automatiques ont été corrigées manuellement pour évaluer la fiabilité de la méthode. Toutefois, pour l'expérience II, les segmentations automatiques ont été utilisées directement dans les analyses (cf. chapitre III 2.2.5).

différentes échelles temporelles et d'examiner leur interdépendance. De plus, grâce aux processus de normalisation, le calcul de la PLV est indépendant de l'amplitude de chaque signal.

Cependant, la méthode de la PLV comporte certaines limitations. D'abord, l'application de cette est limitée aux signaux de forme sinusoïdale. Bien que les courbes d'amplitude puissent être facilement obtenues par modulation d'amplitude, il peut être plus difficile d'obtenir des courbes sinusoïdales pour d'autres propriétés de la parole (ex. la  $f_0$ ).

Par ailleurs, il existe une limite importante de la méthode de la PLV liée au calcul des différences de la phase relative généralisée (cf. chapitre III [3.2.4](#)). Ce calcul nécessite que les valeurs de la phase de deux signaux avec des fréquences différentes soient normalisées par des nombres entiers afin de rendre leurs fréquences similaires. Cette approche repose sur l'hypothèse selon laquelle le ratio de fréquence entre les deux signaux oscillatoires analysés reste constant dans chaque fenêtre d'analyse. Cependant, étant donné que la durée syllabique varie, la fréquence de chaque cycle dans le signal syllabique ne peut pas être constante. Par conséquent, il est possible que le ratio de fréquence entre les signaux (c'est-à-dire le signal de la modulation d'amplitude syllabique et le signal de la modulation d'amplitude supra-syllabique) ne soit pas constant dans chaque fenêtre d'analyse. De plus, plus la fenêtre d'analyse est étroite (ex. dans le cas d'une fenêtre correspondant à un SA où un seul cycle supra-syllabique est détecté dans chaque fenêtre), plus l'influence de cet artefact dû au ratio non constant est importante.

Cependant, cette limitation de la méthode de la PLV ne devrait pas avoir un impact significatif sur les résultats obtenus lors du calcul de la PLV avec une fenêtre couvrant une phrase entière, étant donné que les phrases analysées présentent des structures syntaxiques suffisamment régulières et des placements de proéminence prévisibles. Cette régularité atténue l'impact de la non-constance de la fréquence syllabique.

Dans ce chapitre, nous avons également examiné la relation entre la PLV et le degré de groupement syllabique. Pour rappel, la PLV quantifie la stabilité de la relation temporelle entre deux signaux, à savoir le signal syllabique et le signal supra-syllabique. Nous nous attendions à ce qu'une relation temporelle stable entre la composante syllabique et la composante supra-syllabique corresponde à un degré du groupement syllabique plus prononcé (cf. Figure [5](#)). Cependant, les résultats des analyses présentées dans ce chapitre (cf. [3.2](#)) indiquent que la PLV ne peut quantifier le degré de groupement syllabique que lorsque les deux composantes ne sont pas excessivement similaires, c'est-à-dire lorsque chaque syllabe ne présente pas un niveau de proéminence similaire.

Il est donc essentiel d'interpréter les résultats de la PLV en fonction des caractéristiques spécifiques des données analysées (c'est-à-dire en fonction du ratio entre les fréquences des signaux observés).

---

# CHAPITRE VI : ÉVALUATION DU SENS D'AGENTIVITE ET SON ROLE ORGANISATEUR DANS LE CONTROLE SENSORIMOTEUR

## Sommaire du chapitre

<b>1. Différence individuelle en réponse au décalage continu de la f0 du retour auditif .....</b>	<b>114</b>
<b>2. Relation entre alignement phonétique et sens d'agentivité .....</b>	<b>119</b>
2.1 Comparer la différence de durée entre la syllabe accentuée et non accentuée dans chaque syntagme accentuel.....	119
2.1.1 Comparaison entre la condition avec DécalF0 et la condition sans DécalF0.....	120
2.1.2 Comparaison entre les locutrices dans la « condition perturbations multiples » et celles dans la « condition uniquement DAF ».....	121
2.1.3 Conclusion de la section .....	124
2.2 Comparer le degré de coordination entre les composantes rythmiques .....	124
2.2.1 Comparaison entre la condition avec DécalF0 et la condition sans DécalF0.....	125
2.2.2 Comparaison entre les locutrices dans la « condition perturbations multiples » et celles dans la « condition uniquement DAF ».....	128
2.2.3 Conclusion de la section .....	130
2.3 Comparer les effets indirects du DAF .....	131
2.3.1 Comparaison de l'effet du DAF sur la f0 des voyelles.....	131
2.3.2 Comparaison de l'effet du DAF sur l'intensité des voyelles .....	132
2.3.3 Conclusion de la section .....	135
<b>3. Discussion du chapitre .....</b>	<b>135</b>
3.1 Alignement phonétique en situation d'altération du retour auditif et d'interaction verbale ...	135
3.2 Réduction du sens d'agentivité sous le décalage continu de la f0 du retour auditif.....	137
3.2.1 Mesurer le sens d'agentivité .....	137
3.2.2 Réduction du sens d'agentivité et interaction entre le sens d'agentivité et la prosodie.....	137
3.2.3 Sens d'agentivité en continuum.....	138

Dans le présent chapitre, nous exposerons les résultats de nos analyses visant à examiner nos troisième et quatrième hypothèses (cf. chapitre III [1.2](#)). À des fins de rappel, ces deux hypothèses se définissent comme suit :

**Hypothèse 3 :** Une perturbation continue de f0 du retour auditif diminue le sens d'agentivité des locuteurs, et cette diminution se traduit par une réduction des réponses des locuteurs au DAF.

**Hypothèse 4 :** Dans la régulation du contrôle sensorimoteur, le rôle de la prosodie et celui du sens d'agentivité interagissent, et cette interaction modifie l'organisation rythmique.

Les analyses présentées dans ce chapitre se fondent sur les données issues des locutrices participant à la « condition uniquement DAF » et à la « condition perturbations multiples » de l'expérience II (cf. chapitre III [2.2](#) et [2.3](#)). Dans un premier temps, dans la section [1](#), nous exposons les analyses concernant l'effet du décalage continu de la  $f_0$  du retour auditif (décalF0) dans le but de vérifier la réplication du phénomène d'alignement phonétique ( $f_0$ ), tel qu'observé dans l'étude de Franken et al. (2022) (cf. chapitre II [3.3.2](#)). Nous mettons l'accent sur les différences individuelles pour déterminer si toutes les participantes ont suivi le décalF0 (c'est-à-dire si elles ont aligné leur  $f_0$  à la  $f_0$  décalée du retour auditif).

Ensuite, dans la section [2](#), nous examinons si cet alignement phonétique est associé à une réduction du sens d'agentivité des locutrices en comparant leurs réponses au DAF dans la condition avec décalF0 par rapport à la condition sans décalF0.

Les effets du DAF que nous comparons englobent l'effet du DAF sur la différence de durée entre la voyelle accentuée et la voyelle non accentuée dans chaque SA (cf. chapitre IV [2.2](#)), la coordination entre les composantes rythmiques (cf. chapitre V [2.2](#)), ainsi que certains effets indirects du DAF, tels que son effet sur l'intensité et la  $f_0$ , qui fourniront des observations complémentaires. Ces comparaisons nous permettront également de vérifier si l'effet de l'accentuation interagit avec l'effet sur le sens d'agentivité et si cette interaction influe sur l'organisation rythmique.

## **1. Différence individuelle en réponse au décalage continu de la $f_0$ du retour auditif**

Pour évaluer l'effet du décalF0 et la différence de l'effet du décalF0 entre les participantes, nous avons procédé à une analyse en conduisant *un modèle linéaire séparément pour chaque participante*. Nous avons exclu les essais avec DAF, car nous nous sommes intéressés à l'effet spécifique du décalF0 sur la  $f_0$  produite par chaque locutrice dans cette analyse. Nous avons pris la  $f_0$  médiane des essais de base comme référence, en excluant les autres essais sans décalF0 dans la session de test dans l'expérience II (« condition perturbation multiples », cf. chapitre III [2.3](#)). De plus, nous

avons exclu les répétitions des phrases produites en position 3, c'est-à-dire les répétitions des phrases produites en dernier dans chaque essai<sup>11</sup>.

Les prédicteurs utilisés dans le modèle maximal incluaient le niveau du décalF0 (mesuré par une différence de contraste avec trois niveaux : base, 1 demi-ton de décalage et 2 demi-tons de décalage, avec niveau de base comme référence), le type de phrase, le type de voyelle (mesuré par une différence de contraste avec deux niveaux : les voyelles non accentuables [V1 et V4] et les voyelles accentuables [V2, V3 et V5], avec les voyelles non accentuables comme référence) ainsi que la double interaction entre les prédicteurs. L'effet du type de phrase a été mesuré par la déviation du contraste moyen. La structure des facteurs aléatoires a été automatiquement réduite dans chaque modèle en suivant une procédure de réduction pas à pas basée sur des tests du chi carré appliqués aux résidus des différents modèles possibles. Les modèles conduits pour chaque locutrice ainsi que les résultats clés sont présentés respectivement dans les Tableaux 9 et 10 dans l'Annexe.

La Figure 28 illustre la tendance de la variation de la f0 médiane en réponse au décalF0 pour chaque participante. Les voyelles accentuables et non accentuables sont différenciées par couleur.

### **Résultats :**

Parmi les 10 locutrices faisant partie du **Groupe positif**, c'est-à-dire celles qui ont été perturbées par un décalF0 positif, **huit** ont significativement ajusté leur f0 produite dans la même direction que celle du décalF0. Parmi ces huit locutrices, une seule n'a suivi cette direction qu'avec un niveau du décalF0 plus élevé (c'est-à-dire deux demi-tons), tandis que deux ont même augmenté leur f0 davantage sur les voyelles accentuables (c'est-à-dire les 2èmes, 3èmes et 5èmes voyelles dans les énoncés). Cependant, **deux** locutrices dans ce groupe ont significativement compensé le décalF0 en ajustant leur f0 produite dans la direction inverse du décalage, dont une n'a compensé significativement que lorsque le niveau du décalage était plus élevé.

Dans le **Groupe négatif**, incluant les 10 locutrices perturbées par le décalF0 négatif, **sept** d'entre elles ont significativement suivi la direction du décalF0. Une locutrice a suivi le décalF0 seulement lorsque le niveau était moins élevé, une autre locutrice l'a suivi uniquement sur les voyelles accentuables et lorsque le niveau du décalage était moins élevé. Deux locutrices ne l'ont suivi que lorsque le niveau du décalage était plus élevé et trois l'ont suivi aux deux niveaux du décalage.

---

<sup>11</sup> Les répétitions des phrases qui ont été produites en position 3, c'est-à-dire la dernière phrase de chaque essai, ont été écartées de l'analyse en raison de l'observation fréquente d'un changement de qualité de la voix, voire même de la survenue d'un phénomène de dévoisement à la fin de ces phrases.

Cependant, **deux** locutrices dans ce groupe ont significativement compensé le décalF0 en adaptant leur f0 produite dans la direction inverse du décalage. **Une** autre locutrice n'a pas significativement changé son f0.

Il apparaît que la direction du décalF0 a un impact distinct sur la production entre les locutrices. Parmi les huit locutrices qui ont suivi le décalF0 *positif*, plus le niveau du décalage était élevé, plus elles ont augmenté leur f0. En revanche, parmi les sept locutrices qui ont suivi le décalF0 *néгатif*, seulement deux ont significativement diminué leur f0 davantage lorsque le niveau du décalage était plus élevé (c'est-à-dire deux demi-tons). Les cinq autres locutrices ont davantage diminué leur f0 lorsque le décalF0 était moins élevé (c'est-à-dire un demi-ton), et deux d'entre elles n'ont même pas significativement diminué leur f0 lorsque le décalF0 était plus élevé.

### **Discussion des résultats :**

Les analyses effectuées dans ce chapitre révèlent que la majorité (15 sur 20) des participantes ont suivi le décalF0, confirmant ainsi les résultats obtenus dans l'étude menée par Franken et ses collègues (2022) (cf. chapitre II [3.3.2](#)). Pour ces locutrices qui ont aligné leur f0 à celle de leur retour auditif altéré, le son perçu (c'est-à-dire le retour auditif avec une f0 décalée) serait traité comme s'il émanait d'une source externe (cf. chapitre II [3.1.2](#)).

Malgré le fait que sept des 10 participantes exposées à des modifications négatives de la f0 du retour auditif aient imité la direction du décalage, pour la plupart d'entre elles, le changement de la f0 produite n'était pas proportionnel à l'ampleur du décalF0. Cette observation rejoint les résultats obtenus par Bradshaw et McGettigan (2021) dans leur étude sur la synchronisation de la parole, où ils ont constaté que pour leurs participantes féminines, l'ampleur de convergence vers une cible haute de la f0 était plus importante que vers une cible basse de la f0, et que cette ampleur de convergence n'a pas pu être augmentée par une cible de f0 encore plus basse. Les auteurs ont formulé l'hypothèse que ce phénomène pourrait être simplement dû aux contraintes physiologiques limitant la capacité des femmes à baisser la hauteur de leur voix (en revanche, augmenter la hauteur de la voix pourrait être plus facile pour les femmes).

L'effet du décalF0 est plus prononcé sur les voyelles accentuables pour seulement une petite proportion des locutrices (quatre sur les 15 ayant convergé leur f0 avec celle du retour auditif manipulé). Cela pourrait suggérer que le contrôle de la hauteur de la voix est généralement moins sensible aux facteurs prosodiques que le contrôle temporel de la parole. Toutefois, pour certaines locutrices, les facteurs prosodiques peuvent jouer un rôle dans ce contrôle.

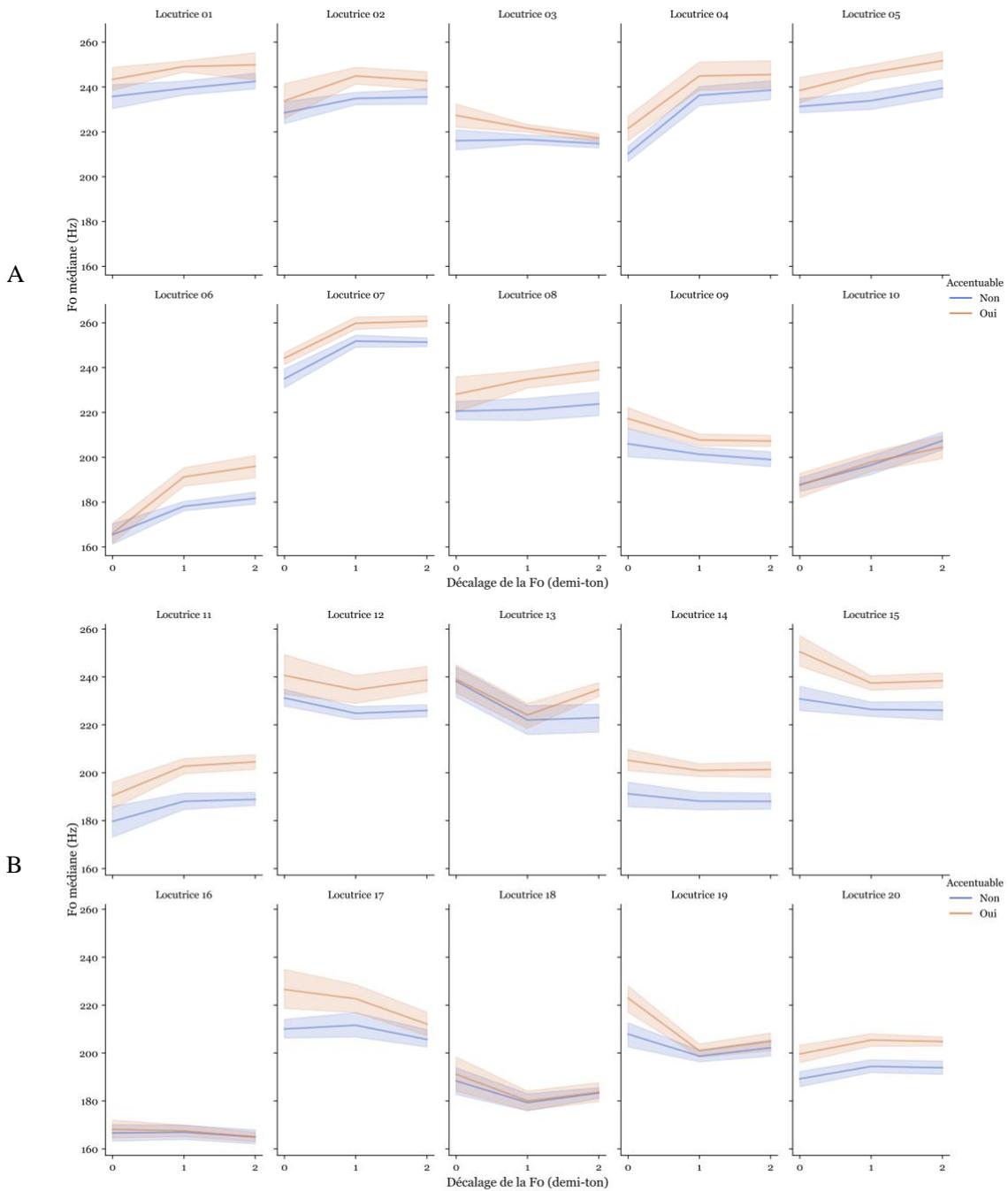
Il convient de souligner que, quelle que soit la direction du décalage  $F_0$ , toutes les locutrices n'ont pas convergé leur  $f_0$  vers celle de leur retour auditif manipulé. Sur les 20 locutrices, quatre ont compensé le décalage de la  $f_0$  du retour auditif en ajustant leur  $f_0$  dans la direction opposée du décalage. En d'autres termes, selon notre interprétation, ces locutrices auraient toujours perçu le son entendu comme étant d'origine interne (cf. chapitre II [3.1.2](#)). Elles ont ajusté leur  $f_0$  afin de réduire le décalage entre la  $f_0$  prédite et celle entendue (c'est-à-dire le retour auditif manipulé). Ainsi, dans les mêmes conditions, le sens d'agentivité relatif au son entendu pendant la production de la parole pourrait varier entre les individus.

En outre, l'une des 20 locutrices n'a pas montré de changement significatif dans sa  $f_0$  sous le décalage  $F_0$ , ce qui suggère que toutes les locutrices ne seraient pas également sensibles aux perturbations du retour auditif. Cette variation de la sensibilité aux perturbations du retour auditif pourrait être liée au fait que certaines locutrices sont moins dépendantes de leur retour auditif (et qu'ils s'appuieraient peut-être d'avantage sur le retour somatosensoriel qui fournit des informations complémentaires au retour auditif, comme l'ont proposé Lametti et al. (2012) (cf. chapitre II [1.2.2](#)).

Figure 28 FO médiane des voyelles (accentuables ou non accentuables) produites à différents niveaux du décalFO, respectivement pour chaque locutrice, dans les essais sans DAF dans l'expérience II (condition « perturbations multiples »)

A : Groupe positif ; B : Groupe négatif

0 = les essais de base ; 1 = les essais avec un demi-ton du décalFO ; 2 = les essais avec deux demi-tons du décalFO



## 2. Relation entre alignement phonétique et sens d'agentivité

Les analyses présentées dans la section précédente (section [1](#)) ont démontré que la majorité des participantes (15 sur 20) ont manifesté un alignement phonétique, c'est-à-dire qu'elles ont suivi le décalF0, tandis qu'une minorité (5 sur 20) de participantes ne l'ont pas manifesté. Nous avons formulé l'hypothèse d'une réduction du sens d'agentivité chez les locutrices ayant suivi le décalF0 (voir chapitre III [1.2](#)). Par conséquent, nous allons maintenant analyser si, pour ces locutrices, leurs réponses au DAF ont diminué dans la condition avec décalF0 par rapport à la condition sans décalF0.

Deux types de comparaisons ont été effectués :

**A)** Comparaison au sein de chaque groupe de locutrices (Groupe Positif ou Groupe Négatif) qui étaient soumises à la « condition perturbations multiples » dans l'expérience II (cf. chapitre III [2.3](#)). Plus précisément, l'ampleur des effets du DAF dans les essais avec décalF0 a été comparée à celle dans les essais sans décalF0.

**B)** Comparaison entre la « condition perturbations multiples » et la « condition uniquement DAF » (cf. chapitre III [2.2](#)). Étant donné que les effets des perturbations dans les essais avec décalF0 et les essais sans décalF0 pour une même locutrice pourraient interagir (en raison de la randomisation des essais expérimentaux dans différentes conditions), nous avons également comparé l'ampleur des effets du DAF entre les locutrices ayant subi le décalF0 (dans les essais avec le décalF0) et celles qui n'ont été perturbées que par le DAF.

### 2.1 Comparer la différence de durée entre la syllabe accentuée et non accentuée dans chaque syntagme accentuel

La comparaison a été effectuée dans le SA initial, englobant les première (V1) et deuxième voyelles (V2) de chaque phrase, ainsi que dans le SA final, composé des deux dernières voyelles de chaque phrase (V4, V5). Il convient de noter que la voyelle accentuée dans chaque segment accentuel est respectivement la voyelle V2, puis la voyelle V5 (cf. chapitre III [3.1](#)).

### 2.1.1 Comparaison entre la condition avec DécalF0 et la condition sans DécalF0

Nous avons d'abord comparé l'ampleur de l'effet du DAF entre les essais avec le décalF0 et ceux sans le décalF0 au sein de chaque groupe de locutrices (Groupe Positif ou Groupe Négatif) qui étaient soumises à la « condition perturbations multiples » dans l'expérience II.

Nous avons conduit des modèles linéaires mixtes distincts pour chacun des deux SAs (SA initial et SA final) dans les trois types de phrases, pour les deux groupes de locutrices ayant subi respectivement un décalF0 positif ou négatif, soit un total de *12 modèles*. Les prédicteurs inclus dans chaque modèle étaient le niveau du décalF0 (niveaux : 0, 1, 2 demi-tons, référence : 0 demi-ton), le niveau du DAF (niveaux : 0, 60, 120 ms, référence : 0 ms, sans condition de base) et le type de voyelle (référence : voyelle non accentuée dans le SA). Des doubles et triples interactions des prédicteurs ont également été incluses. Les données concernant les voyelles présentes dans les SAs finaux des phrases produites en dernière position de chaque essai ont été exclues de l'analyse, en raison de l'atténuation de l'effet du DAF à cet emplacement (cf. chapitre IV [2.2](#)). Les résultats principaux des modèles sont synthétisés dans le Tableau [11](#) dans l'Annexe. Le nombre d'observations incluses dans chaque modèle est indiqué dans la colonne « Phrase-SA » du Tableau. Notre méthode de construction de ces modèles nous permet de comparer directement la différence de l'effet du DAF sur la voyelle accentuée et la voyelle non accentuée dans chaque SA, sans avoir à calculer le ratio entre ces deux types de voyelles.

#### **Résultats :**

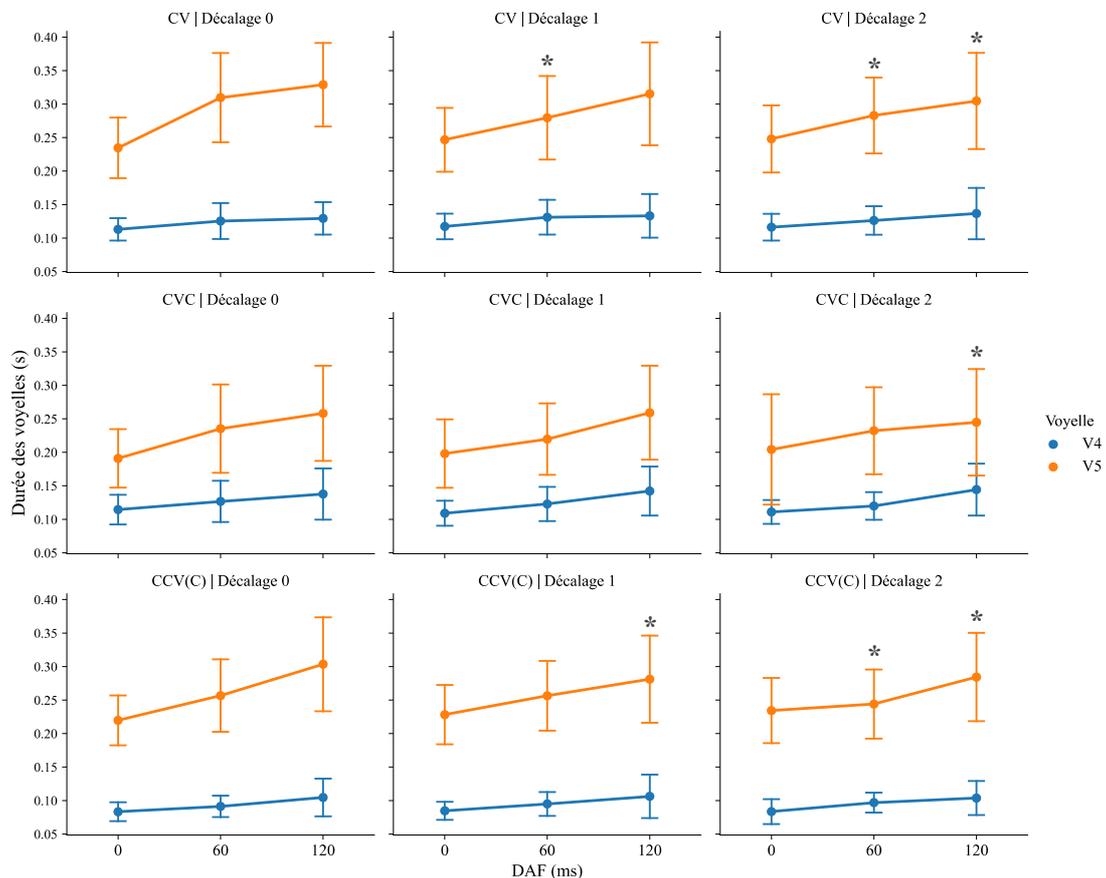
Les voyelles accentuées dans les SAs sont généralement plus allongées par rapport aux voyelles non accentuées sous l'effet du DAF (au moins pour le niveau plus élevé du DAF, soit 120 ms), ce qui est cohérent avec les observations présentées dans le chapitre IV [2.2](#).

Cependant, notre attention se porte principalement sur l'influence du décalF0 sur l'ampleur de l'effet du DAF. Les résultats révèlent que cette influence dépend à la fois de la position des voyelles et de la direction du décalF0. En général, l'effet du décalF0 négatif est très peu prononcé.

Néanmoins, le décalF0 positif présente un effet sur l'effet du DAF dans le SA final. La Figure [29](#) présente la durée des voyelles dans le **SA final** (la voyelle non accentuée V4 et la voyelle accentuée V5) pour chaque niveau du DAF et chaque niveau du décalF0 pour les locutrices du **Groupe positif**. Dans cette condition, la différence d'ampleur de l'effet d'allongement du DAF entre les voyelles accentuées et non accentuées a tendance à **diminuer** dans le **SA final** (cf. les « \* » dans la figure) par rapport à la condition sans décalF0 (cf. la première colonne dans la figure).

Figure 29 Moyenne et écart-type de la durée des voyelles dans le SA final pour chaque niveau du DAF et chaque niveau du **décalF0 positif** dans l'expérience II.

Décalage 0, 1, 2 correspondent respectivement aux essais sans décalage, avec 1 demi-ton de décalage et avec deux demi-tons de décalage. Le symbole \* indique une réduction significative de la différence de l'effet du DAF sur la durée de la voyelle accentuée (courbe en orange) et de la voyelle non accentuée (courbe en bleu) par rapport à la condition sans décalage (première colonne).



### **Conclusion :**

Ces analyses révèlent une tendance à la réduction de la différence dans l'effet d'allongement induit par le DAF entre les voyelles accentuées et les voyelles non accentuées. Cependant, cet effet du décalF0 est très limité (il est observé seulement lorsque le décalF0 est positif et seulement dans le SA final). Il est possible que les effets des perturbations interfèrent entre les essais expérimentaux.

#### **2.1.2 Comparaison entre les locutrices dans la « condition perturbations multiples » et celles dans la « condition uniquement DAF »**

Nous avons conduit six modèles linéaires mixtes distincts, soit un modèle pour chacun des deux SAs dans les trois types de phrases. Les prédicteurs comprenaient le groupe de locutrices (référence : groupe des locutrices dans la « condition uniquement DAF »), le niveau du DAF

(référence : 0 ms, sans condition de base), le type de voyelle (référence : la voyelle non accentuée dans le SA), ainsi que des interactions doubles et triples des prédicteurs. Les données relatives aux voyelles apparaissant dans le SA final des phrases produites dans la dernière position de chaque essai ont été exclues de l'analyse. Les résultats principaux des modèles sont résumés dans le Tableau [12](#) dans l'Annexe. Le nombre d'observations incluses dans chaque modèle est indiqué dans la colonne « Phrase-SA » du Tableau.

### **Résultats :**

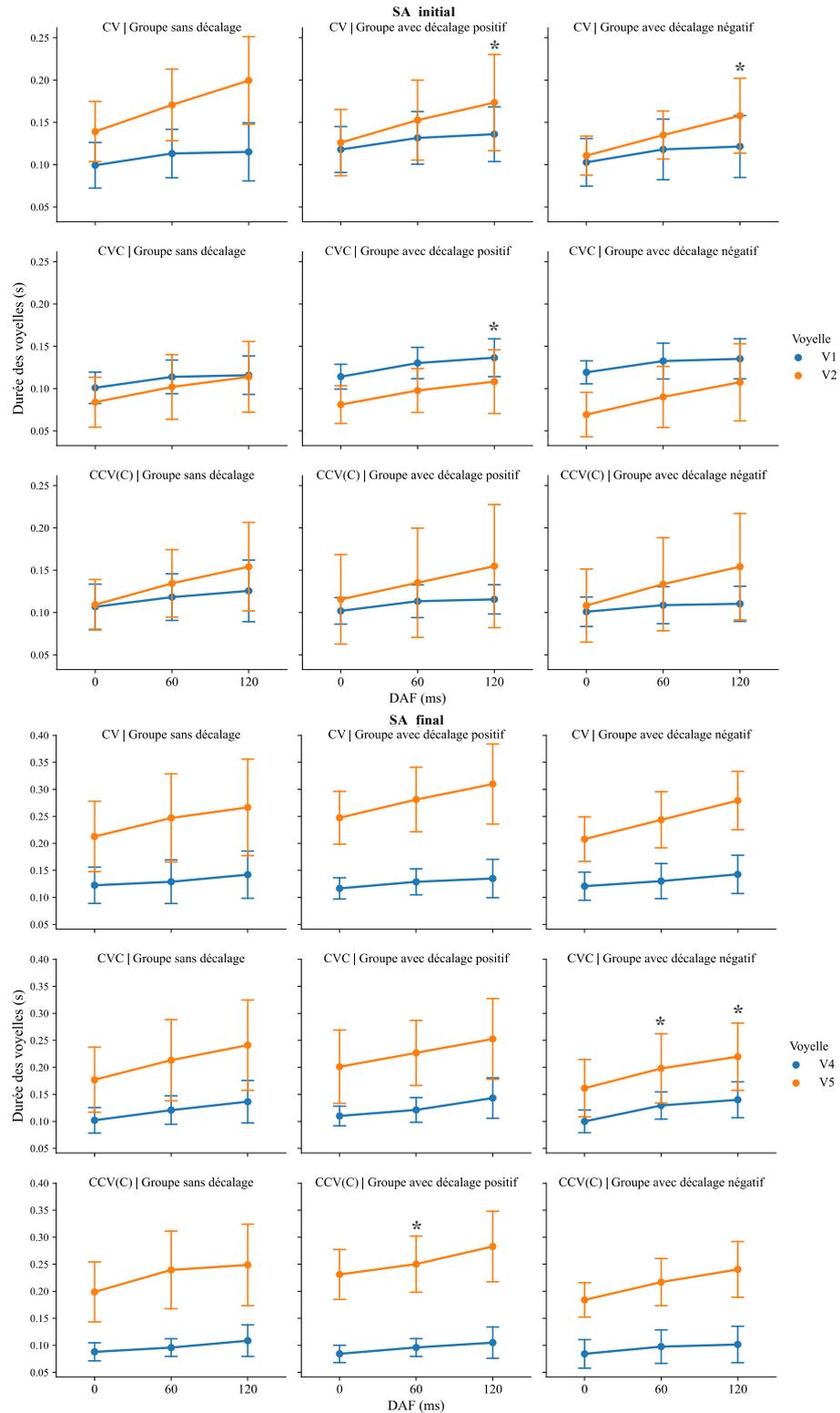
Dans chaque type de phrase, les voyelles accentuées sont plus allongées que les voyelles non accentuées dans les deux SAs. Toutefois, l'effet du groupe de locutrices, qui fait référence à l'effet du décalF0, n'est pas observé dans tous les types de phrases.

La Figure [30](#) présente la durée des voyelles dans le SA initial (en haut) et le SA final (en bas) pour chaque niveau du DAF (axe X) et chaque groupe de locutrices (colonnes), séparément pour chaque type de phrase (lignes). Dans *le SA initial* de la phrase **CV**, nous observons une **réduction significative** de la différence de l'effet d'allongement du DAF à 120 ms sur la voyelle accentuée et la voyelle non accentuée dans le Groupe **positif** et le Groupe **négatif** par rapport au groupe « uniquement DAF ». De même, pour le SA initial de la phrase **CVC**, cette différence s'est avérée **plus petite** dans le Groupe **positif** par rapport au groupe « uniquement DAF ». Cependant, aucune différence significative n'est observée entre les groupes dans le SA initial de la phrase **CCV(C)**.

Dans *le SA final*, nous observons une variation de l'effet du groupe en fonction du type de phrase. Plus précisément, pour la phrase **CV**, aucun effet de groupe n'a été constaté. En revanche, dans la phrase **CVC**, une **réduction significative** de la différence de l'effet d'allongement des deux niveaux du DAF (60 ms et 120 ms) sur la voyelle accentuée et la voyelle non accentuée est observée dans le Groupe **négatif** par rapport au groupe « uniquement DAF ». Quant à la phrase **CCV(C)**, la différence de l'effet du DAF à 60 ms est **plus petite** dans le Groupe **positif**.

Figure 30 Moyenne et écart-type de la durée des voyelles dans chaque SA pour chaque niveau du DAF et chaque groupe de locutrices, dans l'expérience II

Le symbole \* indique une réduction significative de la différence de l'effet du DAF (à niveau marqué) sur la durée de la voyelle accentuée (courbe orange) et de la voyelle non accentuée (courbe bleue) par rapport au groupe sans décalage (première colonne).



### **Conclusion :**

Lors de la comparaison de l'effet du DAF entre les différents groupes de locutrices (Groupe positif ou négatif vs. Groupe « uniquement DAF »), il est observé que tant un décalF0 positif que négatif peuvent réduire l'effet du DAF. De plus, cet effet du décalF0 tend à se manifester dans chaque SA.

#### **2.1.3 Conclusion de la section**

Les analyses présentées dans les sections [2.1.1](#) et [2.1.2](#) mettent en évidence une tendance générale de réduction de l'effet du DAF dans la condition avec décalF0, qu'il soit positif ou négatif, chez les locutrices qui ont suivi le décalF0 (c'est-à-dire celles qui ont aligné leur f0 à la f0 décalée du retour auditif). Cette réduction concerne la différence de durée entre la voyelle accentuée et la voyelle non accentuée dans chaque SA sous l'effet du DAF. En d'autres termes, le rôle organisateur du sens d'agentivité et de l'accentuation interagissent dans le contrôle de l'adaptation au DAF. Cependant, cette réduction n'est pas systématiquement présente dans tous les SAs de tous les types de phrases, ce qui pourrait être partiellement lié à l'instabilité du sens d'agentivité sous l'influence de multiples perturbations du retour auditif.

## **2.2 Comparer le degré de coordination entre les composantes rythmiques**

Les résultats des analyses présentés dans la section précédente ([2.1](#)) révèlent une réduction de la différence entre l'effet du DAF sur la durée de la voyelle accentuée et son effet sur la voyelle non accentuée dans chaque SA pour les locutrices ayant suivi le décalF0. Étant donné que le DAF entraîne également une réorganisation rythmique, renforçant la coordination entre la composante syllabique et la composante supra-syllabique (mesurée par la PLV, cf. chapitre V [2.1](#)), nous nous attendons également à une réduction de cette réorganisation pour les locutrices ayant suivi le décalF0.

De manière similaire, l'effet du DAF sur la PLV sera comparé entre la condition avec le décalF0 et sans décalF0, ainsi qu'entre les locutrices ayant subi le décalF0 et celles de la « condition uniquement DAF ».

De plus, il est opportun de rappeler que nous avons observé une augmentation de la PLV au fil de la production d'un essai. Par conséquent, les répétitions de phrases en positions 2 et 3 des essais présentaient une PLV plus élevée que celles en position 1, même en l'absence du DAF (cf. chapitre V [2.1](#)). Cela a conduit à une neutralisation de l'effet du DAF sur la PLV dans les répétitions de phrases en positions 2 et 3. Nous avons donc décidé de nous concentrer l'influence du décalF0 sur l'effet du DAF sur la PLV dans **répétitions de phrases produites en position 1** dans les essais expérimentaux.

### **2.2.1 Comparaison entre la condition avec DécalF0 et la condition sans DécalF0**

Nous avons également commencé par comparer l'ampleur de l'effet du DAF entre les essais avec le décalF0 et ceux sans décalF0 au sein de chaque groupe de locutrices (Groupe Positif ou Groupe Négatif) qui étaient soumises à la « condition perturbations multiples » dans l'expérience II.

Nous avons conduit des modèles linéaires mixtes distincts pour la PLV des phrases entières et la PLV de chacun des deux SAs (SA initial et SA final)<sup>12</sup>, pour les deux groupes de locutrices ayant respectivement subi un décalF0 positif et un décalF0 négatif, soit un total de *six modèles*. Les prédicteurs inclus dans chaque modèle étaient le niveau du décalF0 (référence : 0 demi-ton), le niveau du DAF (référence : 0 ms, sans condition de base) et le type de phrase. Des doubles et triples interactions des prédicteurs ont également été incluses. L'effet du type de phrase a été codé par la déviation du contraste moyen. Les résultats principaux de ces modèles sont synthétisés dans le Tableau [13](#) dans l'Annexe. Les modèles du groupe exposé aux perturbations positives comportaient un total de 720 observations, tandis que ceux du groupe exposé aux perturbations négatives en avaient 630 chacun.

#### **Résultats du Groupe positif :**

Les résultats montrent que, dans le Groupe positif, le DAF et le décalF0 n'ont pas d'effet significatif sur la PLV des phrases entières et du SA final.

Cependant, les effets du DAF et du décalF0 sont présents lorsque la PLV est analysée au niveau SAs. La Figure [31](#) illustre la PLV du SA initial (en haut) et du SA final (en bas) pour chaque niveau

---

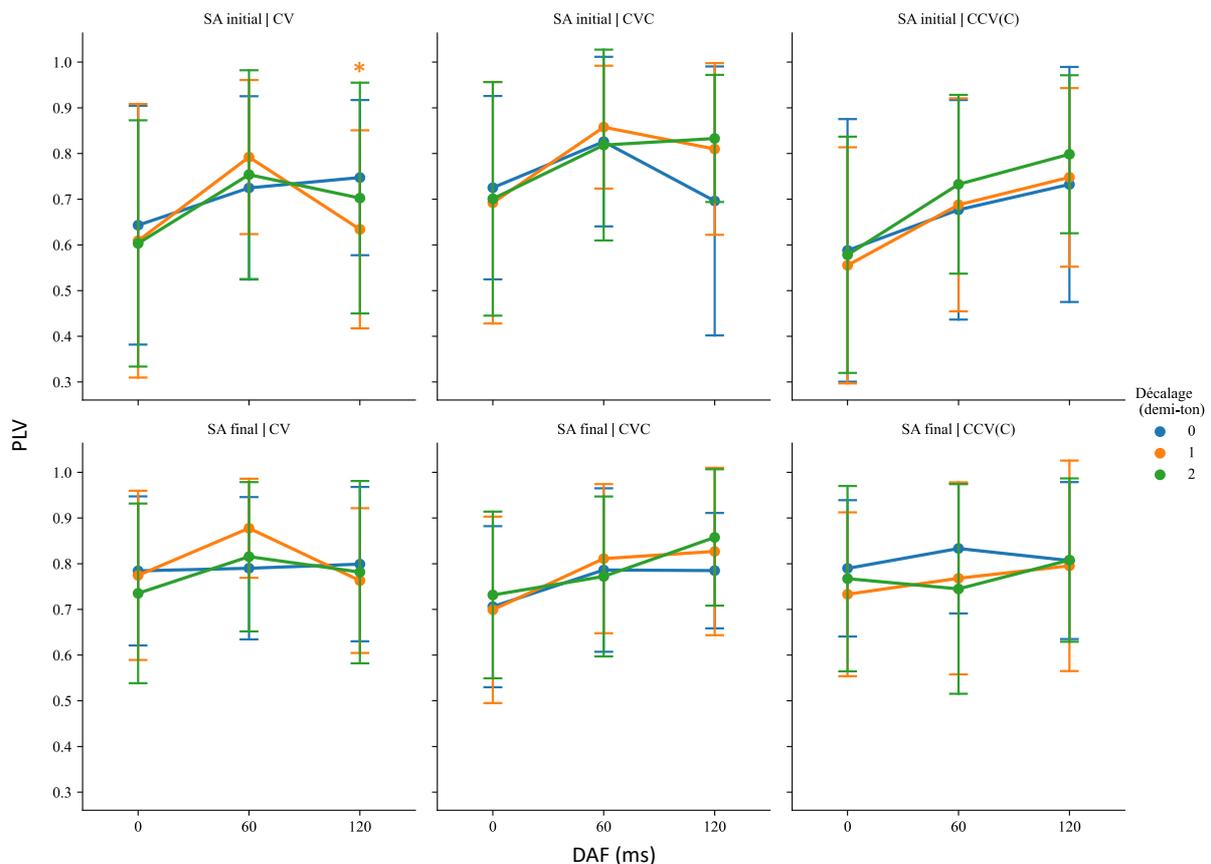
<sup>12</sup> Conformément aux analyses exposées dans la section [2.2](#) du chapitre V, la PLV du SA initial correspond à la moyenne de la PLV dans la portion s'étendant du début de l'énoncé jusqu'à la fin de la deuxième voyelle, tandis que la PLV du SA final correspond à la moyenne de la PLV dans la portion allant du début de la quatrième voyelle jusqu'à la fin de l'énoncé.

du DAF (axe X) et chaque niveau du décalF0 (distingué par la couleur) chez les locutrices ayant suivi décalF0 positif. Le DAF à 60 ms augmente la PLV du **SA initial** et le DAF à 120 ms augmente presque significativement la PLV du **SA initial** dans les essais sans décalF0 (cf. les courbes en bleu dans la figure).

La présence du décalF0 ne diminue pas l'effet moyen du DAF sur la PLV. Cependant, elle a **réduit** l'effet du DAF à 120 ms sur la PLV du **SA initial** dans la **phrase CV** (cf. « \* » en orange dans la figure). Il est à noter que le DAF à 120 ms a tendance à augmenter la PLV du SA initial dans la phrase CV plus que la moyenne des trois types de phrases dans la condition sans décalF0. Bien que cette tendance ne soit pas statistiquement significative, elle pourrait potentiellement expliquer pourquoi la présence du décalF0 ne diminue de manière significative que la PLV du SA initial dans la phrase CV.

Figure 31 PLV du **SA initial** et du **SA final** dans les répétitions de chaque type de phrase en position 1 dans un essai, pour chaque niveau du DAF (axe X) et chaque niveau du décalF0 (couleur), pour les locutrices ayant suivi le décalF0 **positif**, dans l'expérience II

« \* » indique une réduction significative de l'effet du DAF au niveau du décalF0 (indiqué par la couleur) par rapport à la condition sans décalF0 (courbe en bleu).



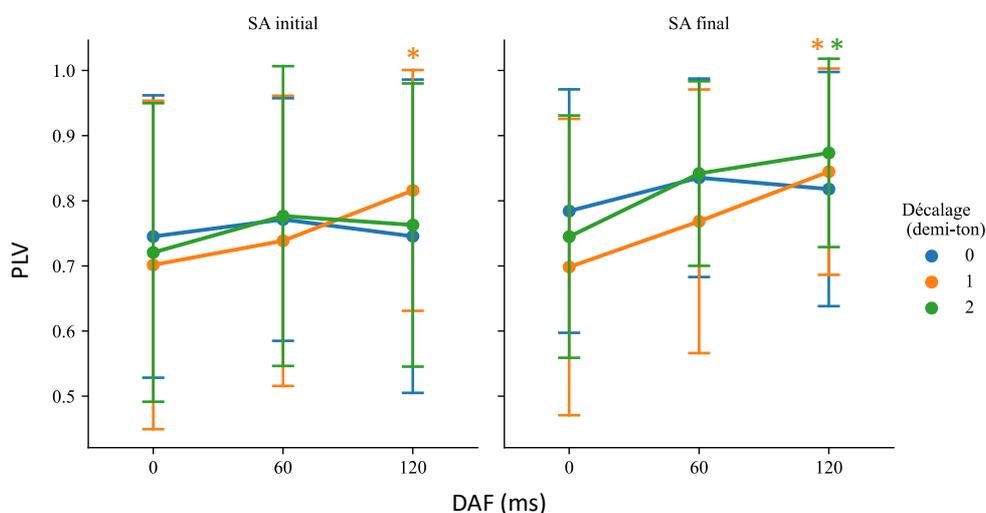
### **Résultats du Groupe négatif :**

Au sein du Groupe négatif, les résultats montrent aucun effet significatif du DAF et du décalF0 sur la PLV des phrases entières.

Les effets du DAF et du décalF0 sont présents lorsque la PLV est analysée au niveau SAs. La Figure 32 illustre la PLV du SA initial (gauche) et du SA final (droite) pour chaque niveau du DAF (axe X) et chaque niveau du décalage de F0 (distingué par la couleur) chez les locutrices ayant suivi le décalF0 négatif. La présence du DAF n'affecte pas de manière significative la PLV du SA initial et du SA final dans les essais sans décalF0 (cf. les courbes en bleu dans la figure). Cependant, lorsque le décalF0 **négatif** est d'un demi-ton, le DAF à 120 ms **augmente** de manière significative la PLV du **SA initial**, et lorsque le décalF0 **négatif** est d'un ou deux demi-tons, il **augmente** de manière significative la PLV du **SA final** (cf. les « \* » en orange et en vert dans la figure).

*Figure 32 PLV du SA initial et du SA final dans les répétitions des phrases en position 1 dans un essai, pour chaque niveau du DAF (axe X) et chaque niveau du décalF0 (couleur), pour les locutrices ayant suivi le décalF0 négatif, dans l'expérience II*

« \* » indique un effet significatif du DAF dans la condition du décalF0 indiquée par la couleur.



### **Conclusion :**

Il semble que l'effet du DAF sur la PLV soit moins prononcé chez les locutrices qui ont suivi le décalF0, même en l'absence de ce dernier. Les raisons de cette réduction de l'effet du DAF ne sont pas bien élucidées. Une possibilité réside dans le fait que, dans la condition sans décalF0, la PLV est déjà relativement élevée (cf. les courbes bleues dans les Figures 31 et 32), et qu'elle n'a pas besoin d'augmenter davantage pour maintenir la stabilité dans le contexte du DAF. Cette observation s'accorde avec nos constatations exposées dans la section 2.2 du chapitre V, selon

lesquelles une fois que la stabilité est atteinte, c'est-à-dire un certain niveau de PLV, il est moins nécessaire d'augmenter la PLV même sous l'effet du DAF.

Une autre possibilité est que la présence du décalF0 ait induit une réduction générale du sens d'agentivité des locutrices (et que cette réduction se serait étendue aux essais sans décalF0). Cependant, cette hypothèse semble moins convaincante, car dans certains cas, la présence du décalF0 a même provoqué une augmentation de l'effet du DAF. Nous supposons que cette augmentation pourrait être liée à l'effet de réduction de la PLV induit par la présence du décalF0. Par exemple, dans le SA final en présence du décalF0 négatif (voir Figure 32, panneau à droite), la PLV est particulièrement basse, c'est-à-dire que la coordination entre les composantes rythmiques est moins stable, dans la condition sans DAF. Il est possible que l'augmentation plus marquée de la PLV soit nécessaire pour renforcer la stabilité. Cette hypothèse concorde également avec nos observations exposées dans la section 2.2 du chapitre V, indiquant qu'un certain niveau de PLV est nécessaire pour maintenir la stabilité dans le contexte du DAF.

Bien que les observations discutées ci-dessus ne permettent pas d'établir un lien entre la présence du décalF0 et la réduction du sens d'agentivité à partir de l'analyse de la PLV, nous constatons tout de même que dans les cas où l'effet du DAF est relativement fort (par exemple, dans le SA initial de la phrase CV), la présence du décalF0 positif réduit l'effet du DAF sur la PLV. Ainsi, l'instabilité du sens d'agentivité demeure une explication plausible dans le cas du décalF0.

### **2.2.2 Comparaison entre les locutrices dans la « condition perturbations multiples » et celles dans la « condition uniquement DAF »**

Trois modèles linéaires mixtes distincts ont été conduits, soit un modèle pour la PLV des phrases entières et un modèle pour chacun des deux SAs. Les prédicteurs incluaient le groupe de locutrices (référence : groupe des locutrices dans la « condition uniquement DAF »), le niveau du DAF (référence : 0 ms, sans condition de base), le type de phrase, ainsi que des interactions doubles et triples des prédicteurs. L'effet du type de phrase a été codé par la déviation du contraste moyen. Les résultats principaux des modèles sont résumés dans le Tableau 14 dans l'Annexe. Chaque modèle a inclus un total de 1974 observations.

#### **Résultats :**

Les résultats ne révèlent **pas de différence significative** de l'effet du DAF sur la PLV des *phrases entières* entre les groupes de locutrices.

Néanmoins, la différence de l'effet du DAF entre les groupes est observée lorsque l'analyse de la PLV est effectuée au niveau des SAs. La Figure 33 présente la PLV du SA initial (en haut) et du SA final (en bas) pour chaque niveau du DAF (axe X) et chaque groupe de locutrices (couleur). Dans *le SA initial*, les deux niveaux du DAF augmentent significativement la PLV dans le SA initial pour les locutrices « uniquement DAF » (cf. les courbes en vert dans la figure). Cependant, l'effet du DAF à 60 ms sur la PLV de la phrase **CCV(C)** est **plus faible** dans le Groupe **négatif** que dans le groupe « uniquement DAF » (cf. le « \* » en orange dans la figure), tandis que l'effet du DAF à 120 ms sur la PLV de la phrase **CV** est **plus faible** dans le Groupe **positif** que dans le groupe « uniquement DAF » (cf. le « \* » en bleu dans la figure).

Il convient de noter que dans les essais sans DAF, la PLV du SA initial de la phrase CCV(C) est significativement plus élevée dans le Groupe négatif que dans le groupe des locutrices « uniquement DAF », ce qui pourrait expliquer la réduction de l'effet du DAF à 60 ms sur la PLV du SA initial de cette phrase dans le Groupe négatif.

Les deux niveaux du DAF augmentent également la PLV dans *le SA final* pour le groupe « uniquement DAF ». Cependant, l'effet du DAF à 120 ms sur la PLV de la phrase **CV** est également **plus faible** dans le Groupe **Positif** que dans le groupe « uniquement DAF » (cf. le « \* » en bleu dans la figure).

### **Conclusion :**

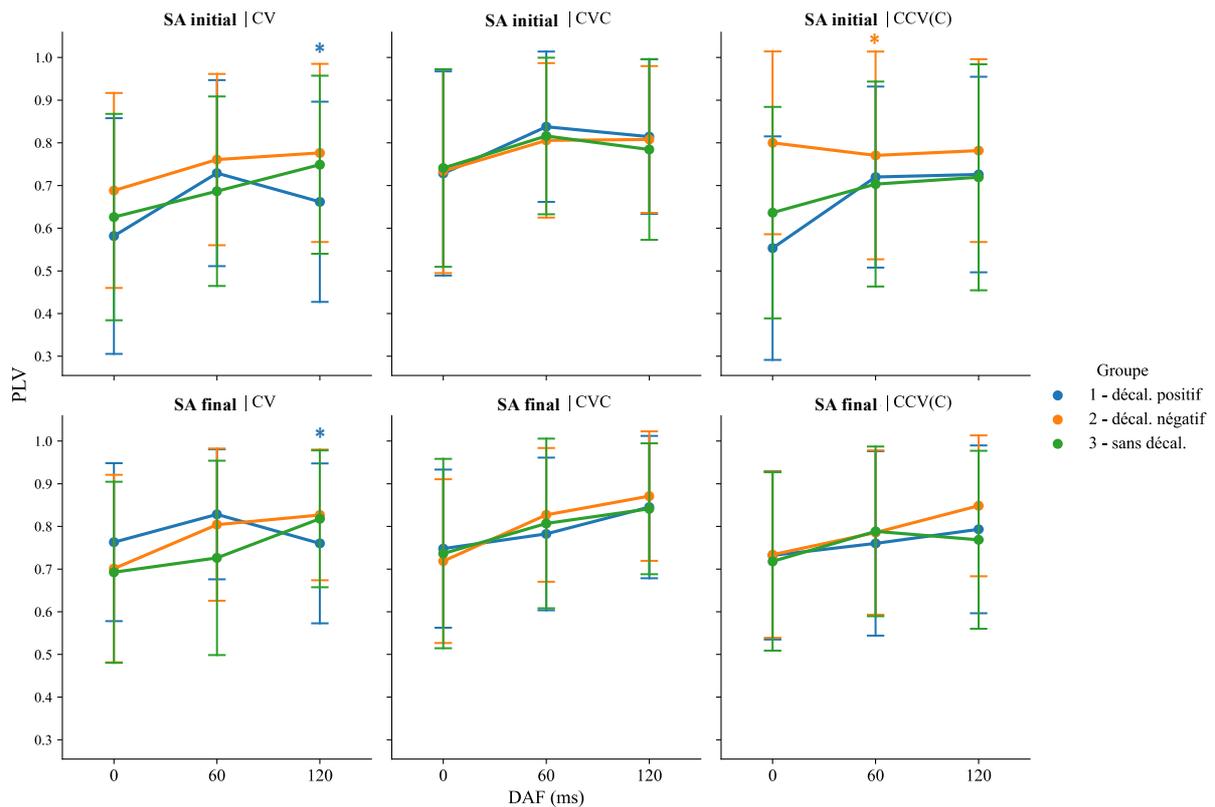
La réduction de l'effet du DAF à 60 ms sur le SA initial de la phrase CCV(C) dans le Groupe négatif est probablement due à une PLV initialement plus élevée pour la phrase CCV(C) de ce groupe en l'absence de DAF par rapport au groupe « uniquement DAF ». Par conséquent, l'effet du DAF sur la PLV du SA initial de la phrase CCV(C) dans le Groupe négatif pourrait être neutralisé. En revanche, l'effet du DAF à 120 ms sur la phrase CV, moins marqué dans le Groupe positif, pourrait être attribué à la présence du décalF0 positif dans ce groupe, étant donné que la PLV de la phrase CV n'est pas significativement différente entre les groupes. Cette observation de l'effet moindre du DAF à 120 ms sur la phrase CV est cohérente avec les résultats de l'analyse présentée dans la section 2.2.1, où il a été démontré que, dans le SA initial de la phrase CV, la présence du décalF0 positif réduit l'effet du DAF.

En comparant l'effet moyen du DAF, aucune différence significative n'a été observée entre les groupes. Cette constatation pourrait être liée au phénomène de renforcement de l'effet du DAF induit par le décalF0, tel que nous l'avons présenté dans la section 2.2.1, car dans cette partie de

l'analyse, nous avons comparé les essais avec décalF0 du Groupe positif ou négatif aux essais du Groupe « uniquement DAF ».

Figure 33 PLV du SA initial et du SA final dans les répétitions des phrases en position 1 dans un essai, pour chaque niveau du DAF (axe X) et chaque groupe de locutrices (couleur), dans l'expérience II

Groupe 1 (bleu) = Groupe Positif; Groupe 2 (orange) = Groupe Négatif; Groupe 3 (vert) = Groupe « uniquement DAF »  
 « \* » indique une réduction significative de l'effet du DAF dans le Groupe (indiqué par la couleur) par rapport au Groupe « uniquement DAF » (bleu).



### 2.2.3 Conclusion de la section

Le rôle du décalF0 dans l'effet du DAF sur la PLV est complexe pour les locutrices ayant suivi le décalF0. D'une part, la présence du décalF0 a tendance à réduire la PLV. L'origine de cette réduction n'est pas claire, mais elle induit une augmentation de l'effet du DAF sur la PLV par rapport à la condition sans décalF0, car un certain niveau de PLV est nécessaire pour assurer la stabilité de la parole dans le contexte du DAF. D'autre part, dans le cas où l'effet du DAF est relativement marqué (l'effet du DAF sur la PLV du SA initial de la phrase CV), cet effet peut également être réduit par la présence du décalF0. Quoi qu'il en soit, une instabilité du sens d'agentivité pourrait être causée par la présence du décalF0.

## 2.3 Comparer les effets indirects du DAF

Les effets du DAF comparés dans les sections [2.1](#) et [2.2](#) sont plutôt les effets directs du DAF, car le DAF est une perturbation temporelle. Cependant, la présence du DAF peut également entraîner des changements sur les autres aspects acoustiques, tels qu'une augmentation de l'intensité et une réduction de la variabilité de la  $f_0$  (cf. chapitre II [1.2.1](#)). Dans cette section, nous présenterons donc les comparaisons des effets du DAF sur l'intensité et la  $f_0$  entre les locutrices dans la « conditions perturbations multiples » et celles dans la « condition uniquement DAF ».

### 2.3.1 Comparaison de l'effet du DAF sur la $f_0$ des voyelles

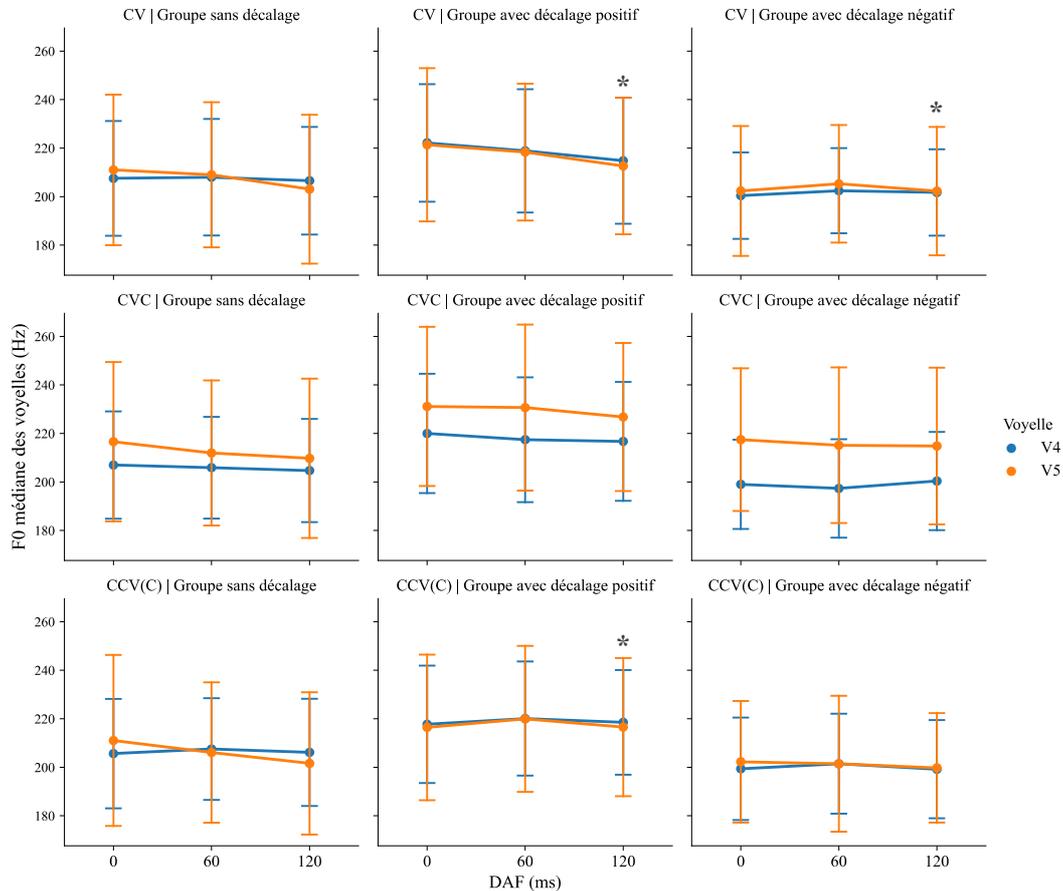
Nous avons conduit *six modèles linéaires mixtes* distincts, soit un modèle pour chacun des deux SAs dans les trois types de phrases. Les prédicteurs comprenaient le groupe de locutrices (référence : Groupe « uniquement DAF »), le niveau du DAF (référence : 0 ms, sans condition de base), le type de voyelle (référence : voyelle non accentuée dans le SA), ainsi que des interactions doubles et triples des prédicteurs. Les données relatives aux voyelles apparaissant dans les SAs finaux des phrases produits dans la dernière position de chaque essai ont été exclues de l'analyse. Les résultats principaux des modèles sont résumés dans le Tableau [15](#) dans l'Annexe. Le nombre d'observations incluses dans chaque modèle est indiqué dans la colonne « Phrase-SA » du Tableau.

#### **Résultats :**

Dans le SA initial, aucun effet des perturbations du retour auditif n'est observé. Cependant, dans le SA final, les effets du DAF et du décalF0 sont observés.

La Figure [34](#) illustre la  $f_0$  médiane des voyelles du **SA final** dans chaque type de phrase (ligne) pour chaque niveau du DAF (axe X) et chaque groupe de locutrices (colonne). Dans ce SA, le DAF (à 60 ou 120 ms) a tendance à réduire la  $f_0$  de la voyelle accentuée (V5) des phrases **CV** et **CCV(C)** pour le groupe « uniquement DAF » (cf. les courbes en orange dans les premier et dernier panneaux de la première colonne dans la figure). Néanmoins, la réduction due au DAF à 120 ms dans la phrase **CV** est **atténuée** dans les Groupes **positif** et **négatif**, de même que celle dans la phrase **CCV(C)** est également atténuée dans le Groupe **positif** (cf. les « \* » dans la figure). En d'autres termes, la  $f_0$  des voyelles accentuées est moins réduite par le DAF à 120 ms, voire n'est pas affectée par le DAF dans le Groupe positif ou négatif.

Figure 34 Moyenne et écart-type de la  $f_0$  médiane des voyelles dans le SA **final** dans chaque type de phrase (lignes) pour chaque niveau du DAF (axe X) et chaque groupe de locutrices (colonne), dans l'expérience II. Le symbole \* indique une réduction significative de l'effet du DAF sur la V5 (courbe orange) par rapport aux locutrices dans la « condition uniquement DAF » (première colonne).



### 2.3.2 Comparaison de l'effet du DAF sur l'intensité des voyelles

Les valeurs d'intensité ont été normalisées en suivant la formule suivante :

$$\hat{I}(t) = \frac{I(t) - I_{min}}{I_{max} - I_{min}}$$

où  $I(t)$  est intensité à l'instant  $t$ .  $I_{min}$  et  $I_{max}$  sont respectivement l'intensité minimale et maximale du signal dans chaque enregistrement. Après cette procédure de normalisation, la valeur d'intensité à chaque instant  $t$  ( $\hat{I}(t)$ ) varie de 0 à 1.

Six modèles linéaires mixtes ont été conduits de la même manière que les modèles conduits pour examiner la différence de l'effet du DAF sur la  $f_0$  des voyelles entre les différents groupes. Les résultats principaux des modèles sont résumés dans le Tableau 16 dans l'Annexe. Le nombre d'observations incluses dans chaque modèle est indiqué dans la colonne « Phrase-SA » du Tableau.

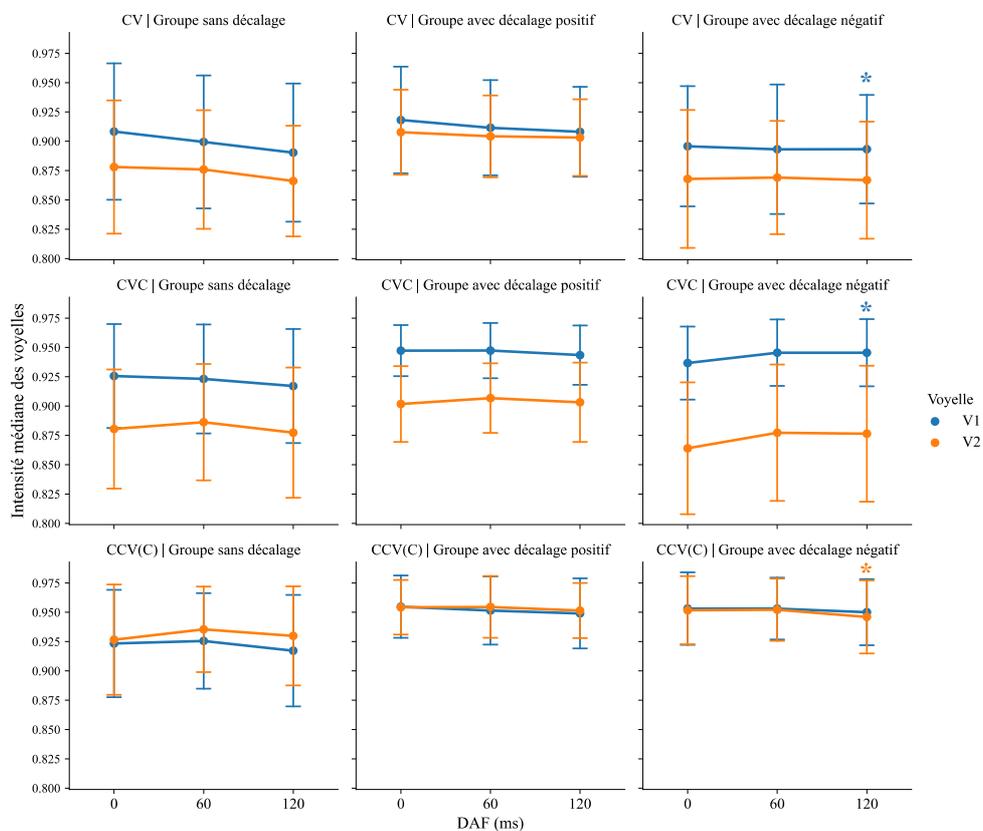
**Résultats pour le SA initial :**

Les Figures 35 présentent la moyenne et l'écart-type de l'intensité médiane des voyelles dans **SA initial** dans chaque type de phrase (ligne) pour chaque niveau du DAF (axe X) et chaque groupe de locutrices (colonne). Dans ce SA, le DAF (à 60 ou 120 ms) a tendance à réduire l'intensité de la voyelle non accentuée (V1) des phrases CV et CVC pour le groupe « uniquement DAF » (cf. les courbes en bleu dans les deux premiers panneaux de la première colonne dans la figure). Néanmoins, la réduction due au DAF à 120 ms dans les phrases **CV** et **CVC** est **atténuée** dans le Groupe **négatif** (cf. les « \* » en bleu dans la figure).

De plus, le DAF (à 60 ou 120 ms) a tendance à augmenter l'intensité de la voyelle accentuée (V2) de la phrase CCV(C) (cf. la courbe en orange dans le dernier panneau de la première colonne dans la figure). Cependant, l'augmentation due au DAF à 120 ms dans la phrase **CCV(C)** est **atténuée** dans le Groupe **négatif** (cf. le « \* » en orange dans la figure).

Figure 35 Moyenne et écart-type de l'intensité médiane des voyelles dans le **SA initial** de chaque type de phrase (ligne) pour chaque niveau du DAF (axe X) et chaque groupe de locutrices (colonne), dans l'expérience II

Le symbole \* indique une réduction significative de l'effet du DAF sur l'intensité médiane de la voyelle (indiquée par la couleur) par rapport à celui dans le Groupe « uniquement DAF » (première colonne).

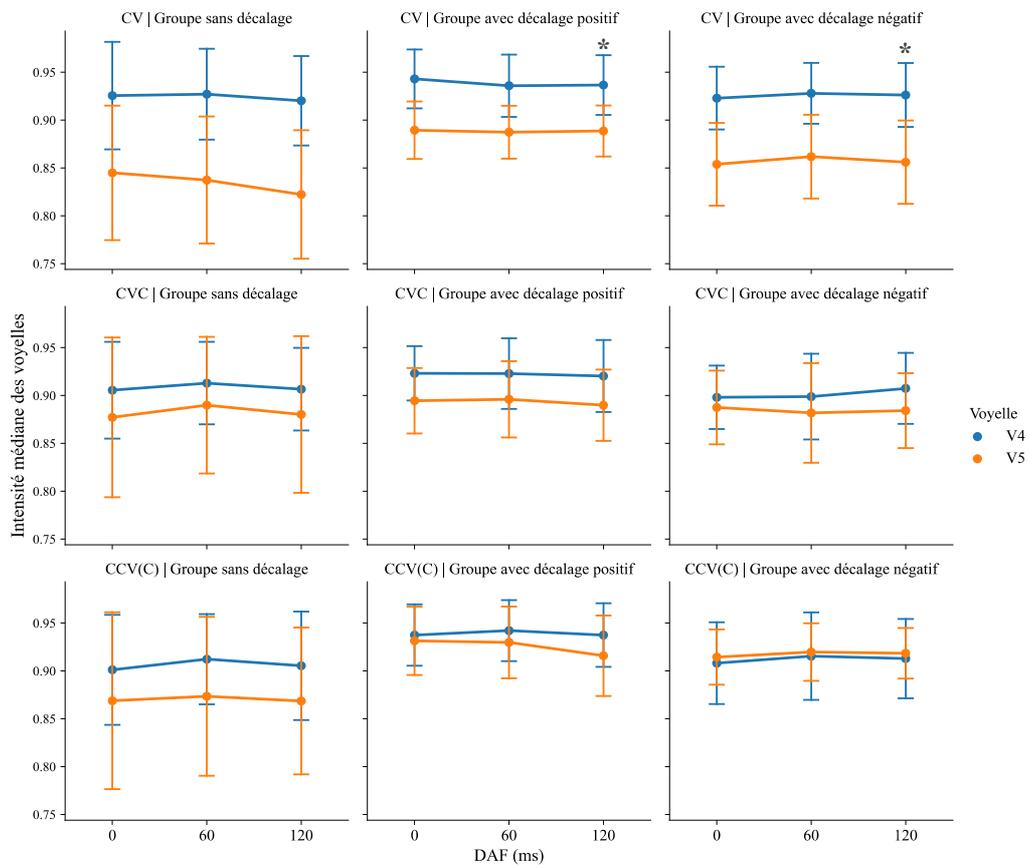


**Résultats pour le SA finale :**

Les Figures 36 présentent la moyenne et l'écart-type de l'intensité médiane des voyelles dans **SA final** dans chaque type de phrase (ligne) pour chaque niveau du DAF (axe X) et chaque groupe de locutrices (colonne). Dans ce SA de la phrase CV, l'intensité de la voyelle accentuée (V5) est moins élevée que celle de la voyelle non accentuée (V4) dans le groupe « uniquement DAF » (cf. le premier panneau de la première colonne dans la figure). Cette différence entre la voyelle accentuée et la voyelle non accentuée est encore renforcée par la présence du DAF (à 60 ou 120 ms). Cependant, cet effet du DAF n'est pas observé dans le Groupe **positif** ou **négatif** (cf. les « \* » dans la figure). En d'autres termes, la présence du DAF dans le Groupe positif ou négatif ne renforce pas la différence entre la voyelle accentuée et la voyelle non accentuée dans le SA final.

Figure 36 Moyenne et écart-type de l'intensité médiane des voyelles dans le **SA final** de chaque type de phrase (ligne) pour chaque niveau du DAF (axe X) et chaque groupe de locutrices (colonne), dans l'expérience II

La présence du DAF (60 et 120 ms) renforce la différence d'intensité entre la V4 et la V5 de la phrase CV dans le groupe « uniquement DAF » (ligne 1 colonne 1). Le symbole « \* » indique une réduction significative de ce renforcement dans les Groupes Positif ou Négatif par rapport au groupe « uniquement DAF ».



### **2.3.3 Conclusion de la section**

Contrairement à l'effet du DAF sur la durée des voyelles, qui consiste en un allongement de chaque voyelle, l'effet du DAF sur la  $f_0$  et l'intensité des voyelles est plus variable. En fonction de la position des voyelles, l'effet du DAF peut ne pas suivre la même direction.

La réduction de l'effet du DAF induite par la présence du décalF0 ne se limite pas à l'effet du DAF sur la durée des voyelles. Les effets indirects du DAF, c'est-à-dire les effets du DAF sur la  $f_0$  et l'intensité, sont également réduits.

## **3. Discussion du chapitre**

Les analyses présentées dans ce chapitre avaient pour objectif d'examiner nos deux dernières hypothèses. Ces deux hypothèses postulaient qu'un décalage continu de  $f_0$  du retour auditif (décalF0) des locutrices réduirait leur sens de l'agentivité relatif à la voix perçue (la voix endogène modifiée) pendant la production de la parole. Cette réduction du sens d'agentivité entraînerait une diminution de l'effet du DAF appliqué simultanément avec le décalF0. De plus, le rôle organisateur du sens d'agentivité devrait interagir avec celui de la prosodie, ce qui se traduirait par une réduction de l'effet du DAF sur la différence de durée entre la voyelle accentuée et la voyelle non accentuée dans chaque SA, ainsi qu'une réduction de l'effet du DAF sur la PLV (la coordination entre les composantes rythmiques).

Les résultats de ces analyses ont mis en évidence une perte partielle (ou une instabilité) du sens d'agentivité. Dans la suite de cette section, nous discuterons d'abord des différences individuelles liées à l'effet du décalF0 sur la production de la  $f_0$  (section [3.1](#)), puis nous explorerons les implications de la réduction du sens d'agentivité induite par le décalF0 (section [3.2](#)).

### **3.1 Alignement phonétique en situation d'altération du retour auditif et d'interaction verbale**

La majorité des participantes à cette étude ont aligné leur  $f_0$  sur la  $f_0$  altérée de leur retour auditif. Ce phénomène d'alignement phonétique est similaire à celui de la convergence phonétique observé

dans les situations d'interaction verbale (Pardo, 2006 ; Cummins, 2002 ; Bradshaw et McGettigan, 2021). Comme cela a été expliqué dans la section [3.1](#) du chapitre II, l'alignement phonétique pourrait être interprété comme une conséquence de la mise à jour de nos représentations internes de la cible du contrôle sensorimoteur, lesquelles intègrent les caractéristiques de la parole de notre interlocuteur. Par conséquent, les prédictions des énoncés à venir seraient formulées en se basant sur ces représentations mises à jour (Pickering et Garrod, 2013). Il est donc raisonnable de supposer que lorsque la  $f_0$  du retour auditif est continuellement altérée, la plupart des locutrices mettraient également à jour leur cible sensorimotrice. En d'autres termes, ces locuteurs ont tendance à traiter leur propre voix (altérée) comme si elle était générée par une source externe.

Il convient de souligner que, dans les deux conditions étudiées en littérature (le décalF0 et l'interaction verbale), certaines locutrices et certains locuteurs ne se sont pas alignés sur la voix perçue (c'est-à-dire leur propre voix altérée ou la voix de leur interlocuteur). Nos résultats ont montré qu'une petite partie des participantes (soit 5 sur 20) n'ont pas suivi le décalF0. De même, certaines études sur l'interaction verbale ont également révélé qu'une faible proportion de participants n'ont pas imité leur interlocuteur (ex. Bradshaw et al., 2021).

Dans notre étude, parmi les cinq locutrices qui n'ont pas suivi le décalF0, une seule n'a ni significativement suivi ni compensé le décalF0. Il est possible de supposer que cette locutrice est moins susceptible aux perturbations du retour auditif. En revanche, les quatre autres participantes ont compensé de manière significative le décalF0. Ainsi, ces quatre locutrices ne semblent pas être moins sensibles aux perturbations que les locutrices ayant suivi décalF0 du retour auditif. L'une des hypothèses envisageables est que ces locutrices attribueraient davantage de confiance à leur retour somatosensorielle ou à leur mécanisme prédictif (cf. chapitre II [1.1](#)), qui est influencé par leur expérience passée. De la même manière, il est raisonnable de supposer que les locuteurs qui ne cherchent pas à imiter leur interlocuteur lors des interactions verbales pourraient accorder moins d'importance aux informations provenant de leur interlocuteur dans leur contrôle sensorimoteur.

## **3.2 Réduction du sens d'agentivité sous le décalage continu de la f0 du retour auditif**

### **3.2.1 Mesurer le sens d'agentivité**

Franken et al. (2021) ont tenté de mesurer le sens d'agentivité des locuteurs ayant été soumis à un décalage continu de grande amplitude de la f0 dans leur retour auditif en observant leur réaction à des décalages courts et imprévisibles de faible amplitude de la f0 dans le retour auditif (cf. chapitre II [3.4](#)). Comme nous l'avions déjà proposé dans la section 3.4 du chapitre II, le décalage continu de la f0 dans le retour auditif des locuteurs pourrait provoquer un basculement de la cible sensorimotrice. En d'autres termes, lors des processus de contrôle sensorimoteur, la f0 décalée de faible ampleur serait comparée à un niveau de f0 ajusté en raison du décalage de grande ampleur de la f0. Toute divergence découlant de cette comparaison pourrait déclencher une réponse compensatrice afin de s'aligner sur ce nouveau niveau de f0.

Cependant, il est important de noter que le basculement de la cible sensorimotrice est souvent associé à la perception de la voix d'une autre personne au cours d'une interaction verbale (cf. chapitre II [3.1.2](#)). Ce basculement amène les locuteurs à imiter la voix de cette autre personne. Dans l'étude menée par Franken et al. (2021), les participants ont eu tendance à imiter la f0 décalée lorsque le décalage était plus important, se comportant comme si la voix qu'ils entendaient était celle d'une autre personne. Cette observation est cohérente avec les résultats de notre étude.

Dans le cadre de cette thèse, nous avons choisi de mesurer le sens d'agentivité en évaluant l'ampleur de l'effet du DAF plutôt que de comparer les effets obtenus avec différents degrés de décalage de la f0 dans le retour auditif. En effet, le DAF induit une perturbation de la dimension temporelle de la parole, qui ne serait pas comparée à la cible de la f0 dans le contrôle sensorimoteur. Les résultats de ce chapitre ont montré une réduction de l'effet du DAF chez les locutrices ayant aligné leur f0 à la f0 décalée de leur retour auditif, ce qui suggère une réduction partielle de leur sens d'agentivité. Nous examinerons plus en détail cette réduction dans la section suivante ([3.2.2](#)).

### **3.2.2 Réduction du sens d'agentivité et interaction entre le sens d'agentivité et la prosodie**

Pour les locutrices ayant aligné leur f0 à la f0 décalée de leur retour auditif, une réduction a été observée dans tous les aspects des effets du DAF présentés dans cette thèse. Plus précisément, chez ces locutrices, la présence du décalF0 a eu tendance à diminuer la différence de l'effet

d'allongement du DAF entre la voyelle accentuée et la voyelle non accentuée dans chaque SA (cf. [2.1](#)). Elle a également réduit l'effet du DAF sur la coordination entre les composantes rythmiques (PLV) dans certaines conditions (cf. [2.2](#)). De plus, elle a même réduit les effets indirects du DAF sur la  $f_0$  et l'intensité (cf. [2.3](#)).

Il convient de noter que la réduction de tous les aspects des effets du DAF n'était pas systématiquement présente dans toutes les conditions (par exemple, en fonction du Groupe Positif ou Négatif, ou des types de phrases). Cette observation suggérerait une certaine instabilité du sens d'agentivité des locutrices en raison de la présence du décalF0, car il est difficile d'établir une règle précise pour déterminer la présence ou l'absence de la réduction des effets du DAF en fonction des différentes conditions.

Nos résultats suggèrent également une interaction entre le rôle organisateur de la prosodie et celui du sens d'agentivité dans les processus du contrôle sensorimoteur. Sous l'effet du décalF0, il est très rare d'observer une réduction de l'effet d'allongement du DAF sur les voyelles non accentuées. La réduction majeure a été constatée dans la différence de cet effet entre la voyelle accentuée et la voyelle non accentuée au sein de chaque SA. En d'autres termes, le rôle de l'accentuation dans le contrôle sensorimoteur dans la condition du DAF a été atténué en présence du décalF0. De manière similaire, la réduction de l'effet du DAF sur la PLV, observée dans certaines conditions, suggère qu'il y a eu un affaiblissement du renforcement de la coordination entre la composante syllabique et la composante supra-syllabique dans la condition du DAF lorsque le décalF0 était présent, ce qui semble indiquer une adaptation de la structure coordinative du rythme en réponse à la présence du décalF0.

### **3.2.3 Sens d'agentivité en continuum**

Les résultats de notre étude mettent en évidence une perte partielle du sens d'agentivité chez les locutrices lorsque la  $f_0$  de retour auditif est continuellement décalée. Cette observation suggère que la génération du sens d'agentivité, tout comme d'autres processus cognitifs tels que la perception des phonèmes, ne peut être considérée comme un processus catégoriel. En d'autres termes, le sens d'agentivité relatif à une action ne varie pas simplement d'une catégorie (ex. « je suis l'auteur de l'action ») à une autre (ex. « je ne suis pas l'auteur de l'action »). Au contraire, le sens d'agentivité varie en continuum en fonction de la situation dans laquelle l'action est produite. Par exemple, Blakemore et ses collègues (2000) ont mené une étude sur le chatouillement et ont découvert que

celui-ci augmentait progressivement à mesure que la divergence entre le retour sensoriel attendu et le retour sensoriel réel augmentait, suite à une manipulation artificielle (cf. chapitre II [3.2.2](#)). Nous avons tous l'expérience que le chatouillement est normalement moins intense lorsque nous nous chatouillons nous-mêmes. Cela est dû à l'atténuation sensorielle qui accompagne une action produite par soi-même. Ainsi, l'augmentation progressive de la divergence entre le retour sensoriel attendu et le retour sensoriel réel décrite dans l'étude de Blakemore et al. (2000) a entraîné une perte progressive du sens d'agentivité chez les participants.

Concernant l'étude présentée dans le cadre de cette thèse, il est possible que plusieurs facteurs aient contribué à ce que les participantes aient seulement partiellement perdu leur sens d'agentivité relatif à leur propre voix lorsqu'elles étaient soumises au décalF0.

Dans un premier temps, il convient de noter que la parole est un phénomène multidimensionnel. Dans cette étude particulière, bien que la f0 de la voix des participantes ait été altérée, d'autres dimensions de la parole telles que les formants, sont restées inchangées.

De plus, le sens d'agentivité dépend de l'intégration de signaux pertinents pour des différentes modalités sensorielles (Botvinick et Cohen, 1998). Cependant, en fonction de la situation, certaines modalités sensorielles peuvent avoir un poids plus important que d'autres dans la génération du sens d'agentivité et peuvent même induire un faux sens d'agentivité<sup>13</sup>. Dans notre étude, malgré l'altération continue de la f0 du retour auditif des locutrices, leur retour somatosensoriel demeurait inchangé. Ainsi, le retour somatosensoriel non altéré est resté disponible et a continué à jouer un rôle dans la génération du sens d'agentivité relatif à leur propre voix.

En dernier lieu, il est à noter que malgré les perturbations du retour auditif, les locutrices peuvent toujours compter sur leur mécanisme prédictif (cf. chapitre II [1.1](#)), lequel s'appuie partiellement sur les expériences antérieures pour effectuer des prédictions. Même si les informations auditives fournies par le retour auditif peuvent être moins fiables, le mécanisme prédictif pourrait aider les locutrices à maintenir partiellement leur sens d'agentivité relatif à leur propre voix.

---

<sup>13</sup> Par exemple, l'expérience de l'illusion de la main en caoutchouc démontre que les signaux visuels seuls peuvent suffire à produire un sens d'agentivité chez la majorité des participants (Botvinick et Cohen, 1998).



---

# CHAPITRE VII : CONCLUSION GENERALE ET PERSPECTIVES

## 1. Conclusion

L'objectif principal de cette thèse est d'explorer l'interaction entre le rôle organisateur de la prosodie et celui du sens de l'agentivité dans le contrôle sensorimoteur lors de la production de la parole. Cette recherche s'inscrit dans une perspective intégrative visant à examiner comment la complexité et la flexibilité interagissent dans le processus de production de la parole. Pour ce faire, notre travail commence par une revue de la complexité inhérente au contrôle moteur lors de la production de la parole, mettant en lumière le rôle du groupement rythmique, qui est modulé par la prosodie, dans la simplification et la stabilisation du contrôle moteur de la parole (Chapitre [I](#)). Ensuite, nous avons examiné le mécanisme du contrôle sensorimoteur qui permet au système moteur de s'adapter de manière flexible à différentes situations de production de la parole, ainsi que le sens de l'agentivité relatif à notre propre voix, résultant des processus de contrôle sensorimoteur et influençant la manière dont le système sensorimoteur traite notre voix par rapport à celle d'une autre personne (Chapitre [II](#)). Cette base théorique établie nous a permis de formuler des questions de recherche, de développer les hypothèses y sont associées et de définir nos méthodes pour les analyses (Chapitre [III](#)). Enfin, nous avons présenté les résultats de nos analyses visant à examiner chacune de nos hypothèses (Chapitres [IV](#), [V](#), [VI](#)). Nous procéderons à un résumé plus approfondi de chaque étape de cette thèse.

Dans le **chapitre [I](#)**, nous avons tout d'abord examiné une source significative de la complexité dans le processus de production de la parole, à savoir la nécessité de coordonner

plusieurs articulateurs pour atteindre les objectifs linguistiques. Cependant, cette coordination génère un rythme coordinatif (un rythme collectif aux articulateurs) qui contribue à réduire la complexité du contrôle moteur et, par conséquent, à stabiliser la coordination des articulateurs (Poeppel et Assaneo, 2020). Ce rythme coordinatif correspond généralement à la production des syllabes (voir Figure 1). Néanmoins, la génération de ce rythme est soumise à diverses contraintes d'ordre biomécanique, neurophysiologique et cognitif (MacNeilage, 1998 ; Davis et MacNeilage, 2004 ; Di Cristo, 2003 ; Assaneo et Poeppel, 2018), et elle est également modulée par les caractéristiques prosodiques spécifiques à chaque langue.

L'accentuation, en tant que paramètre essentiel de la prosodie, a pour effet de mettre en relief certaines syllabes par rapport aux autres au sein des énoncés, créant ainsi des groupements de syllabes en unités délimitées par les accents (Troubetzkoy, 1949). L'accentuation est organisée de manière hiérarchique, créant une structuration des éléments en unités de différentes tailles. Par exemple, en français, l'unité prosodique la plus petite, constituée de plusieurs syllabes et marquée par des accents primaires, est le mot prosodique (Astésano, 2017) ou le syntagme accentuel (Jun et Fougeron, 2000), selon différents auteurs. Deux ou plusieurs syntagmes accentuels peuvent à leur tour former des unités plus larges, les syntagmes intonatifs, délimités par des accents plus proéminents que ceux délimitant les syntagmes accentuels (Jun et Fougeron, 2000).

Le rythme de la parole repose sur le groupement syllabique, lequel est hiérarchiquement organisé par la structure prosodique. Les études visant à caractériser les régularités rythmiques en utilisant des mesures de durée (par exemple, Pike, 1945 ; Ramus et al., 1999 ; Low et al., 2000 ; Grabe et Low, 2002) ont souvent produit des résultats non cohérents. D'une part, ces études n'ont pas tenu compte de la hiérarchisation du rythme. D'autre part, il s'avère pratiquement impossible de prendre en considération tous les facteurs (tels que le débit de parole et le style de parole) qui peuvent influencer la durée des éléments (ex. les syllabes) lors de l'analyse des régularités rythmiques.

Une autre série d'études (Leong et al., 2014 ; Falk et Kello, 2017 ; Lancia et al., 2023) a révélé que le degré et la régularité du groupement syllabique correspondent à une relation

temporelle stable entre la composante syllabique et la composante supra-syllabique (cf. Figure 5). La composante syllabique est associée à la production des syllabes, tandis que la composante supra-syllabique est liée à la production des syllabes proéminentes, lesquelles peuvent être modulées par l'accentuation (mais ne dépendent pas exclusivement de celle-ci). Cette perspective d'analyse des régularités rythmiques considère le rythme comme un mécanisme de coordination visant à stabiliser la production de la parole.

Dans le **chapitre II**, nous avons passé en revue les aspects du contrôle sensorimoteur, qui nous permettent d'ajuster de manière flexible notre production dans différentes conditions. Au sein de ce système, les retours sensoriels sont comparés à nos prédictions des résultats des articulations, ce qui permet une évaluation en temps réel de l'état du système de production de la parole et une adaptation rapide de notre production si nécessaire (Parrell et al., 2019 ; Wolpert et al., 1995 ; Miall et Wolpert, 1996). Les études portant sur la manipulation du retour auditif (Burnett et al., 1998 ; Houde et Jordan, 1998, 2002 ; Larson et al., 2007 ; Jones et Munhall, 2003) ont démontré que lorsque le retour auditif est perturbé (par exemple, par un décalage positif de la  $f_0$  du retour auditif), la plupart des locuteurs effectuent une compensation motrice (par exemple, en réduisant la  $f_0$  produite) pour réduire la divergence entre le retour auditif et leurs prédictions.

La prosodie pourrait constituer un facteur influençant le contrôle sensorimoteur. Par exemple, les voyelles accentuées s'adaptent de manière plus flexible (ex. en prolongeant davantage leur durée) pour faire face aux perturbations temporelles par rapport aux voyelles non accentuées (Kalveram et Jancke, 1989 ; Oschkinat et Hoole, 2022). Cependant, la manière dont le rythme s'organise dans le contexte de cette interaction entre la prosodie et le contrôle moteur n'a pas fait l'objet de nombreuses études.

Le contrôle sensorimoteur est également influencé par le sens de l'agentivité, qui concerne notre perception du contrôle de nos propres actions (Haggard, 2017). Le sens de l'agentivité revêt une importance cruciale, car il nous permet de différencier notre propre voix de celle d'une autre personne, et ces deux voix sont traitées de manière distincte par notre système de contrôle sensorimoteur. Une voix exogène met à jour notre cible sensorimotrice

(Pickering and Garrod, 2004), nous incitant à imiter cette voix (Pardo, 2006 ; Pardo et al., 2013 ; Cummins, 2002 ; Bradshaw et McGettigan, 2021). Le sens de l'agentivité découle de la comparaison entre les retours sensoriels et nos prédictions (Farrer et al., 2008 ; Blakemore et al., 2000). Ainsi, les perturbations du retour auditif pourraient entraîner une diminution du sens de l'agentivité, ce qui, à son tour, pourrait influencer les processus de contrôle sensorimoteur.

Sur la base du cadre théorique établi dans les deux premiers chapitres, le **chapitre III** de notre étude a été consacré à la formulation de nos questions de recherche. Ces questions portent sur l'interaction entre la prosodie et les processus de contrôle sensorimoteur, l'influence de cette interaction sur l'organisation rythmique (la coordination entre les composantes rythmiques), ainsi que la manière dont le sens de l'agentivité participe à cette interaction. Afin d'explorer ces questions de recherche, nous avons conduit deux expériences en utilisant le paradigme du retard du retour auditif (DAF) et le décalage continu de la  $f_0$  du retour auditif (décalF0), impliquant la participation de 36 locutrices natives du français. Notre première hypothèse suggère que les voyelles accentuées s'adapteraient de manière plus flexible au DAF, ce qui se traduirait par un allongement plus marqué sous l'effet du DAF. De plus, cet allongement serait lié à leur position dans la structure prosodique, mais ne serait pas simplement prédit par la durée plus longue des syllabes accentuées. Partant de cette première hypothèse, nous postulons que le rythme, modulé par l'accentuation, serait réorganisé sous l'influence du DAF. Plus précisément, dans le contexte du DAF, la coordination entre la composante syllabique et la composante supra-syllabique serait renforcée pour garantir la stabilité de la production de la parole. En ce qui concerne le sens de l'agentivité, nous nous attendions à ce qu'un décalage continu de la  $f_0$  du retour auditif réduise le sens de l'agentivité des locutrices relatif à leur propre voix, ce qui se traduirait par une atténuation des effets du DAF. De plus, cette réduction affecterait notamment la différence de l'effet d'allongement induit par le DAF sur les voyelles accentuées et non accentuées, ainsi que la réorganisation rythmique provoquée par le DAF.

Nous avons évalué la différence dans l'effet d'allongement du DAF entre les voyelles accentuées et non accentuées en analysant le ratio entre la durée de ces deux types de voyelles. Pour mesurer le degré de coordination entre la composante syllabique et la composante supra-syllabique du rythme de la parole, nous avons utilisé la méthode du *Phase Locking Value* (PLV). Cette méthode repose sur l'analyse de la modulation de l'amplitude du signal acoustique, ce qui nous a permis d'extraire deux signaux oscillatoires. Le premier est associé aux changements d'amplitude à l'échelle temporelle de la production des syllabes, et l'autre aux changements plus lents observés à l'échelle temporelle définie par la modulation de la prééminence syllabique. La PLV est inversement proportionnelle à la variabilité de la relation temporelle entre ces deux signaux oscillatoires.

Dans le **chapitre IV**, les résultats des analyses portant sur la différence dans l'effet d'allongement du DAF sur la durée des voyelles accentuées et non accentuées ont confirmé que les voyelles accentuées présentaient généralement une plus grande flexibilité d'adaptation au DAF, c'est-à-dire qu'elles étaient plus allongées en réponse au DAF. De plus, cette flexibilité ne pouvait pas être simplement attribuée à la durée plus longue des voyelles accentuées, mais était plutôt liée à leur statut prosodique. Cependant, il a également été observé que l'effet d'allongement du DAF interagissait avec l'allongement final. Lorsque les voyelles avaient déjà subi un allongement excessif à la frontière droite d'un syntagme intonatif, le rôle de l'accentuation dans la modulation de l'effet du DAF était annulé. Autrement dit, lorsque les voyelles étaient déjà fortement allongées (ex. à la fin des énoncés), la différence de durée entre les voyelles accentuées et non accentuées ne s'accroissait pas, même en présence d'un DAF à un niveau élevé.

Dans le **chapitre V**, les résultats des analyses portant sur l'effet du DAF sur la PLV ont confirmé que, dans le contexte du DAF, la coordination entre la composante syllabique et la composante supra-syllabique du rythme de la parole a été renforcée. Il est à noter que ce renforcement de la coordination pouvait se produire naturellement à mesure que la production d'un flux de parole progressait, même en l'absence du DAF. Par conséquent,

lorsque le degré de cette coordination était suffisamment élevé pour garantir la stabilité de la production de la parole dans le contexte du DAF, un renforcement supplémentaire de cette coordination devenait superflu. Il est à noter que l'établissement de ce renforcement de la coordination pouvait être un processus graduel, et il s'avérait plus rapide pour les phrases ayant une structure syntaxique simple.

Cependant, nous avons observé deux stratégies différentes de réorganisation rythmique parmi les locutrices. Certaines d'entre elles ont eu tendance à adopter une production de style saccadé dans le contexte du DAF, en accordant une proéminence similaire à chaque syllabe. Cette production de style saccadé entraînait également une PLV élevée en raison de la grande similarité entre la composante syllabique et la composante supra-syllabique. Cependant, pour les locutrices qui avaient recours à cette stratégie, la coordination des deux composantes rythmiques ne semblait pas augmenter avec l'augmentation du DAF. En revanche, pour les autres locutrices qui n'avaient pas opté pour la production de style saccadé, le degré de coordination des deux composantes rythmiques augmentait avec l'augmentation du DAF, et ce renforcement pouvait être associé à un renforcement du groupement syllabique ou à une augmentation de la régularité du groupement syllabique. Par ailleurs, il convient de noter que le groupement physique des syllabes peut être moins efficacement prédit par l'accentuation en français (cf. Figure 12) du fait que l'intensité joue un rôle moins significatif dans le processus d'accentuation.

La réorganisation rythmique sous l'effet du DAF nous suggère que le rythme peut également s'adapter de manière flexible à différentes situations du contrôle sensorimoteur dans la production de la parole. Cette observation offre une perspective intéressante pour l'utilisation de la manipulation du retour auditif dans la conception d'approches thérapeutiques visant à traiter les troubles de la parole liés à des problèmes d'organisation rythmique, tels que le bégaiement, ainsi que dans la conception d'approches visant à faciliter l'acquisition des caractéristiques rythmiques propres à des langues présentant des patrons rythmiques complexes.

Enfin, dans le **chapitre VI**, les résultats d'une première partie d'analyse ont révélé que la plupart des locutrices avaient tendance à suivre un décalage continu de la  $f_0$  dans leur retour auditif. Cela suggère qu'elles ont ajusté leur cible de  $f_0$  dans le contrôle sensorimoteur, comme si elles percevaient leur propre voix modifiée comme celle d'une autre personne. Les résultats d'une deuxième partie d'analyse ont montré que pour ces locutrices qui ont suivi le décalage continu de la  $f_0$  dans leur retour auditif, c'est-à-dire celles qui ont aligné leur  $f_0$  sur la  $f_0$  décalée du retour auditif, les effets du DAF décrits dans les chapitres IV et V, à savoir la différence de durée accrue entre les syllabes accentuées et non accentuées, ainsi que le renforcement de la coordination entre les composantes rythmiques, avaient tendance à diminuer. De plus, même les effets indirects du DAF sur la  $f_0$  et l'intensité étaient réduits chez ces locutrices.

Ces résultats indiquent une réduction du sens d'agentivité chez ces locutrices, et cette réduction a eu un impact sur leur contrôle sensorimoteur, mettant en lumière une interaction entre le sens d'agentivité et l'accentuation dans les processus de contrôle sensorimoteur pendant la production de la parole.

Il convient de noter que la réduction des effets du DAF n'était pas systématiquement présente dans toutes les conditions, ce qui pourrait être attribué à l'instabilité du sens d'agentivité provoquée par le décalage continu de la  $f_0$  du retour auditif. De plus, ces résultats ont montré que le sens d'agentivité varierait sur un continuum, et dans le cadre de cette étude, une perte partielle du sens d'agentivité des locutrices pourrait être due à plusieurs facteurs, tels que la non-altération d'autres dimensions de la voix (par exemple, les formants) et la disponibilité du retour somatosensoriel.

En résumé, cette thèse souligne l'importance d'aborder l'étude de la production de la parole dans une perspective intégrative.

## 2. Perspectives

Cette thèse ouvre la voie à de nombreuses perspectives de recherche.

### 2.1 Rôle de l'accentuation dans le contrôle sensorimoteur pendant la production de la parole

Les résultats des analyses menées dans cette thèse ont mis en évidence le rôle de l'accentuation dans le contrôle sensorimoteur de la production de la parole. Cette thèse s'est principalement concentrée sur les accents primaires, en l'occurrence les accents finaux. Cependant, comme mentionné dans la section [2.2.2](#) du chapitre I, la langue française permet également la présence d'accents initiaux sur les premières syllabes des syntagmes accentuels (Di Cristo, 2003 ; Jun et Fougeron, 2000). Les accents initiaux présentent des caractéristiques distinctes des accents finaux, telles qu'un allongement de l'attaque des syllabes (Jankowski et al., 1999 ; Astésano, 2001). Les accents initiaux tendent à apparaître davantage lorsque deux accents primaires sont séparés par un grand nombre de syllabes non accentuées (Jankowski et al., 1999 ; Jun et Fougeron, 2000). Les accents initiaux contribuent donc à l'équilibre rythmique. Cela suscite un intérêt particulier pour une exploration plus approfondie du rôle des accents initiaux dans le contrôle sensorimoteur de la production de la parole, ainsi que pour la recherche de différences entre le rôle des accents initiaux et celui des accents finaux.

De plus, en raison des limites des stimuli, notre étude n'a pas permis une exploration approfondie du rôle des différents niveaux d'accentuation, qui pourraient correspondre à différents niveaux de proéminence (cf. Figure [4](#) représentant la grille métrique). Au cours de nos analyses, il est apparu que le rôle de l'accentuation à la frontière d'un syntagme intonatif était annulé par un allongement final excessif causé par un ton montant à la fin des phrases (cf. chapitre IV [2.3](#)). Il convient de noter que le français présente divers types de syntagmes intonatifs (Delais-Roussarie, 2022). Par conséquent, il serait pertinent d'approfondir l'étude du rôle des différents niveaux d'accentuation dans les futures recherches.

## 2.2 Organisation rythmique dans le contrôle sensorimoteur pendant la production de la parole

Nous avons observé un renforcement de la coordination entre la composante syllabique et la composante supra-syllabique du rythme de la parole afin d'assurer la stabilité de la production de la parole dans le contexte du DAF chez les locutrices francophones. Toutefois, comme évoqué dans la section 3.2.2 du chapitre I, certaines langues peuvent naturellement présenter un degré de coordination entre les composantes rythmiques plus élevé que d'autres langues. Par exemple, des études ont montré que le degré de cette coordination est plus élevé en allemand qu'en français (Lancia et al., 2019). En réalité, le degré plus élevé de cette coordination en allemand pourrait être lié à une corrélation plus prononcée entre l'accentuation et la prééminence de l'énergie (Delattre, 1966b), ce qui n'est pas systématiquement observé en français (cf. Figure 12). Il serait donc intéressant d'explorer comment l'organisation rythmique évolue dans ces langues dans le contexte des perturbations temporelles telles que le DAF. Par ailleurs, il serait pertinent d'explorer, en ce qui concerne le français, les modulations de la  $f_0$  ainsi que la coordination entre les modulations de la  $f_0$  et de l'amplitude dans le contexte des perturbations.

Dans cette perspective des modulations de la  $f_0$ , une comparaison entre les langues présentant des caractéristiques de macrorythme distinctes serait particulièrement intéressante. Le macrorythme se réfère à la saillance de l'alternance entre des sons graves et aigus (LH-LH-LH ...) ou des sons aigus et graves (HL-HL-HL ...) (Jun, 2012 ; Polyanskaya et al., 2020). Par exemple, le mandarin, en raison de ses tons lexicaux, présente un patron tonal moins régulier que le français. L'impact de cette irrégularité tonale sur l'organisation rythmique dans le contexte de perturbations du retour auditif pourrait ainsi être exploré.

De plus, comme nous l'avons précédemment évoqué dans le chapitre I (cf. 1.2), la coordination des mouvements des articulateurs génère le rythme syllabique. Cependant, comme l'a suggéré Goldstein (2019), les mouvements des articulateurs présentent également une régularité. Par conséquent, il serait pertinent d'explorer l'organisation rythmique des mouvements des articulateurs dans le contexte de perturbations du retour auditif.

Finalement, comme discuté précédemment dans le chapitre V (voir section [4.4](#)), il serait également intéressant d'explorer la mise en place de la manipulation du retour auditif en tant qu'outil pour faciliter l'acquisition du langage ou comme une approche thérapeutique pour les troubles de la parole.

### **2.3 Flexibilité du sens d'agentivité pendant la production de la parole**

Une perte partielle du sens d'agentivité induite par le décalage continu de la  $f_0$  dans le retour auditif nous encourage à approfondir la recherche sur les facteurs qui peuvent influencer le sens d'agentivité pendant la production de la parole.

De plus, il serait pertinent de comparer l'organisation rythmique dans le contexte des perturbations du retour auditif, où le sens d'agentivité des locuteurs relatif à leur propre voix est réduit, avec celle dans le contexte de la coordination entre individus (par exemple, la parole synchronisée), où un sens d'agentivité commun (le sentiment d'agir ensemble) pourrait émerger (Shiraishi et Shimada, 2021).

### **2.4 Modulation de la flexibilité du contrôle sensorimoteur dans la production de la parole**

Jusqu'à présent, seuls quelques modèles ont pris en compte la modulation temporelle de la parole. Par exemple, le modèle neuronal pour la dynamique temporelle de l'intégration sensorimotrice, proposé par Assaneo et Poeppel (2018), permet de décrire l'effet de l'allongement du DAF sur les syllabes grâce à la synchronisation des activités entre le cortex moteur et le cortex auditif. Alors que l'enveloppe de la parole se synchronise avec les activités dans le cortex auditif, la présence du DAF entraîne une diminution de la fréquence fondamentale des oscillations dans les cortex auditif et moteur. Cependant, le rôle de la structure prosodique dans ce modèle demeure peu clair. Selon les résultats présentés dans cette thèse, il est important de prendre en considération les aspects prosodiques qui

influencent le contrôle temporel dans les processus du contrôle sensorimoteur lors de la modélisation de la production de la parole. De plus, il serait également pertinent de tenir compte du sens d'agentivité, qui influe sur l'ampleur de l'adaptation du système moteur aux perturbations, dans la modélisation.



---

## REFERENCES

- Abbs, J. H., et V. L. Gracco. 1984. « Control of Complex Motor Gestures: Orofacial Muscle Responses to Load Perturbations of Lip during Speech ». *Journal of Neurophysiology* 51(4):705-23. doi: [10.1152/jn.1984.51.4.705](https://doi.org/10.1152/jn.1984.51.4.705).
- Abney, Drew H., Alexandra Paxton, Rick Dale, et Christopher T. Kello. 2014. « Complexity Matching in Dyadic Conversation ». *Journal of Experimental Psychology: General* 143(6):2304-15. doi: [10.1037/xge0000021](https://doi.org/10.1037/xge0000021).
- Agnello, Joseph G. 1970. « Durational Differences in Speech Production under Normal and Delayed Auditory Feedback ». *Speech Monographs* 37(3):195-98. doi: [10.1080/03637757009375665](https://doi.org/10.1080/03637757009375665).
- Arvaniti, Amalia. 2009. « Rhythm, Timing and the Timing of Rhythm ». *Phonetica* 66(1-2):46-63. doi: [10.1159/000208930](https://doi.org/10.1159/000208930).
- Arvaniti, Amalia. 2012. « The Usefulness of Metrics in the Quantification of Speech Rhythm ». *Journal of Phonetics* 40(3):351-73. doi: [10.1016/j.wocn.2012.02.003](https://doi.org/10.1016/j.wocn.2012.02.003).
- Assaneo, M. Florencia, et David Poeppel. 2018. « The Coupling between Auditory and Motor Cortices Is Rate-Restricted: Evidence for an Intrinsic Speech-Motor Rhythm ». *Science Advances* 4(2):eaao3842. doi: [10.1126/sciadv.aao3842](https://doi.org/10.1126/sciadv.aao3842).
- Astésano, Corine. 1999. « Rythme et discours: invariance et sources de variabilité des phénomènes accentuels en français ». PhD Thesis, Aix-Marseille 1.
- Astésano, Corine. 2001. *Rythme et accentuation en français : invariance et variabilité stylistique*. Editions L'Harmattan.
- Astésano, Corine. 2017. « Le statut de l'accent initial dans la phonologie prosodique du français : enjeux descriptifs et psycholinguistiques ». Habilitation à diriger des Recherches, UT2J.
- Astésano, Corine, Roxane Bertrand, Robert Espesser, et Noël Nguyen. 2012. « Perception des frontières et des proéminences en français ». P. 353-60 in *Proceedings of the Joint Conference JEP-TALN-RECITAL 2012, volume 1: JEP*.
- Astésano, Corine, Cyrille Magne, Radouane El Yagoubi, et Mireille Besson. 2004. « Influence du rythme sur le traitement sémantique en français: approches comportementale et électrophysiologique ». *Soumis dans ce volume*.
- Atkinson, Chester J. 1953. « Adaptation To Delayed Side-Tone ». *Journal of Speech and Hearing Disorders* 18(4):386-91. doi: [10.1044/jshd.1804.386](https://doi.org/10.1044/jshd.1804.386).
- Aubanel, Vincent, et Noël Nguyen. 2020. « Speaking to a Common Tune: Between-Speaker Convergence in Voice Fundamental Frequency in a Joint Speech Production Task » édité par F. J. Torreira. *PLOS ONE* 15(5):e0232209. doi: [10.1371/journal.pone.0232209](https://doi.org/10.1371/journal.pone.0232209).

- Bachrach, David L. 1964. « Sex Differences in Reactions to Delayed Auditory Feedback ». *Perceptual and Motor Skills* 19(1):81-82. doi: [10.2466/pms.1964.19.1.81](https://doi.org/10.2466/pms.1964.19.1.81).
- Behroozmand, Roozbeh, Laura Karvelis, Hanjun Liu, et Charles R. Larson. 2009. « Vocalization-Induced Enhancement of the Auditory Cortex Responsiveness during Voice F0 Feedback Perturbation ». *Clinical Neurophysiology* 120(7):1303-12. doi: [10.1016/j.clinph.2009.04.022](https://doi.org/10.1016/j.clinph.2009.04.022).
- Behroozmand, Roozbeh, Oleg Korzyukov, Lindsey Sattler, et Charles R. Larson. 2012. « Opposing and Following Vocal Responses to Pitch-Shifted Auditory Feedback: Evidence for Different Mechanisms of Voice Pitch Control ». *The Journal of the Acoustical Society of America* 132(4):2468-77. doi: [10.1121/1.4746984](https://doi.org/10.1121/1.4746984).
- Bertinetto, Pier Marco, et Chiara Bertini. 2008. « On modeling the rhythm of natural languages ». P. 427-30 in *Proceedings of the Fourth International Conference on Speech Prosody*.
- Black, John W. 1951. « The Effect Of Delayed Side-Tone Upon Vocal Rate And Intensity ». *Journal of Speech and Hearing Disorders* 16(1):56-60. doi: [10.1044/jshd.1601.56](https://doi.org/10.1044/jshd.1601.56).
- Blakemore, Sarah-Jayne, Daniel Wolpert, et Chris Frith. 2000. « Why Can't You Tickle Yourself? ». *NeuroReport* 11(11):R11-16. doi: [10.1097/00001756-200008030-00002](https://doi.org/10.1097/00001756-200008030-00002).
- Boersma, Paul. 2001. « Praat, a system for doing phonetics by computer ». *Glott international* 5:9(10):341-45.
- Botvinick, Matthew, et Jonathan Cohen. 1998. « Rubber Hands 'Feel' Touch That Eyes See ». *Nature* 391(6669):756-756. doi: [10.1038/35784](https://doi.org/10.1038/35784).
- Bradshaw, Abigail R., et Carolyn McGettigan. 2021. « Convergence in Voice Fundamental Frequency during Synchronous Speech » édité par V. Kempe. *PLOS ONE* 16(10):e0258747. doi: [10.1371/journal.pone.0258747](https://doi.org/10.1371/journal.pone.0258747).
- Browman, Catherine P., et Louis Goldstein. 1989. « Articulatory Gestures as Phonological Units ». *Phonology* 6(2):201-51. doi: [10.1017/S0952675700001019](https://doi.org/10.1017/S0952675700001019).
- Browman, Catherine P., et Louis Goldstein. 1990. « Tiers in articulatory phonology, with some implications for casual speech ». P. 341-76 in *Papers in Laboratory Phonology*, édité par J. Kingston et M. E. Beckman. Cambridge University Press.
- Browman, Catherine P., et Louis Goldstein. 1992. « Articulatory Phonology: An Overview ». *Phonetica* 49(3-4):155-80. doi: [10.1159/000261913](https://doi.org/10.1159/000261913).
- Burnett, Theresa A., Marcia B. Freedland, Charles R. Larson, et Timothy C. Hain. 1998. « Voice F0 Responses to Manipulations in Pitch Feedback ». *The Journal of the Acoustical Society of America* 103(6):3153-61. doi: [10.1121/1.423073](https://doi.org/10.1121/1.423073).
- Cai, Shanqing, Marc Boucek, Satrajit S. Ghosh, Frank H. Guenther, et Joseph S. Perkell. 2008. « A system for online dynamic perturbation of formant trajectories and results from perturbations of the Mandarin triphthong/iau ». P. 65-68 in *8th International Seminar on Speech Production (ISSP)*. Strasbourg, France.
- Casserly, Elizabeth D. 2011. « Speaker Compensation for Local Perturbation of Fricative Acoustic Feedback ». *The Journal of the Acoustical Society of America* 129(4):2181-90. doi: [10.1121/1.3552883](https://doi.org/10.1121/1.3552883).

- Caudrelier, Tiphaine, et Amélie Rochet-Capellan. 2019. « Changes in speech production in response to formant perturbations: An overview of two decades of research », P. 15-75 in *Speech production and perception: Learning and memory*. Vol. 6, *Speech Production and Perception*, édité par S. Fuchs, J. Cleland, et A. Rochet-Capellan. Peter Lang.
- Chang, Edward F., Caroline A. Niziolek, Robert T. Knight, Srikantan S. Nagarajan, et John F. Houde. 2013. « Human Cortical Sensorimotor Network Underlying Feedback Control of Vocal Pitch ». *Proceedings of the National Academy of Sciences* 110(7):2653-58. doi: [10.1073/pnas.1216827110](https://doi.org/10.1073/pnas.1216827110).
- Chartier, Josh, Gopala K. Anumanchipalli, Keith Johnson, et Edward F. Chang. 2018. « Encoding of Articulatory Kinematic Trajectories in Human Speech Sensorimotor Cortex ». *Neuron* 98(5):1042-1054.e4. doi: [10.1016/j.neuron.2018.04.031](https://doi.org/10.1016/j.neuron.2018.04.031).
- Chesters, Jennifer, Ladan Baghai-Ravary, et Riikka Möttönen. 2015. « The Effects of Delayed Auditory and Visual Feedback on Speech Production ». *The Journal of the Acoustical Society of America* 137(2):873-83. doi: [10.1121/1.4906266](https://doi.org/10.1121/1.4906266).
- Christoffels, Ingrid K., Elia Formisano, et Niels O. Schiller. 2007. « Neural Correlates of Verbal Feedback Processing: An fMRI Study Employing Overt Speech ». *Human Brain Mapping* 28(9):868-79. doi: [10.1002/hbm.20315](https://doi.org/10.1002/hbm.20315).
- Cichocki, Wladyslaw. 2015. « The timing of Accentual Phrases in read and spontaneous speech: data from Acadian French ». *Canadian Acoustics* 43(3).
- de Cornulier, Benoît. 1995. « Métrique ». *Encyclopédia Universalis*.
- Cummins, Fred. 2002. « On Synchronous Speech ». *Acoustics Research Letters Online* 3(1):7-11. doi: [10.1121/1.1416672](https://doi.org/10.1121/1.1416672).
- Cummins, Fred, et Robert Port. 1998. « Rhythmic Constraints on Stress Timing in English ». *Journal of Phonetics* 26(2):145-71. doi: [10.1006/jpho.1998.0070](https://doi.org/10.1006/jpho.1998.0070).
- Curio, Gabriel, Georg Neuloh, Jussi Numminen, Veikko Jousmäki, et Riitta Hari. 2000. « Speaking Modifies Voice-Evoked Activity in the Human Auditory Cortex ». *Human Brain Mapping* 9(4):183-91. doi: [10.1002/\(SICI\)1097-0193\(200004\)9:4<183::AID-HBM1>3.0.CO;2-Z](https://doi.org/10.1002/(SICI)1097-0193(200004)9:4<183::AID-HBM1>3.0.CO;2-Z).
- Dauer, R. M. 1983. « Stress-Timing and Syllable-Timing Reanalyzed ». *Journal of Phonetics* 11(1):51-62. doi: [10.1016/S0095-4470\(19\)30776-4](https://doi.org/10.1016/S0095-4470(19)30776-4).
- Dauer, Rebecca M. 1987. « Phonetic and phonological components of language rhythm ». P. 447-50 in *Proceedings of the 11th international congress of phonetic sciences*. Vol. 5.
- Davis, Barbara L., et Peter F. MacNeilage. 2004. « The frame/content theory of speech evolution: From lip smacks to syllables ». *Primatologie* 6:305-28.
- De Jong, Kenneth, Mary E. Beckman, et Jan Edwards. 1993. « The Interplay Between Prosodic Structure and Coarticulation ». *Language and Speech* 36(2-3):197-212. doi: [10.1177/002383099303600305](https://doi.org/10.1177/002383099303600305).
- Delais-Roussarie, Elisabeth. 1993. « Structure rythmique et prosodique du français: pour une approche psycho-cognitive ». *Linx* 29(2):45-71.

- Delais-Roussarie, Elisabeth. 2022. « Prosodic Structure Revisited: The Need to Disentangle Rhythm from Intonation ». *Romanistisches Jahrbuch* 73(1):31-69. doi: [10.1515/roja-2022-0002](https://doi.org/10.1515/roja-2022-0002).
- Delais-Roussarie, Elisabeth, et Brechtje Post. 2008. « Unités prosodiques et grammaire de l'intonation: vers une nouvelle approche ». *Actes des Journées d'étude sur la Parole JEP-TALN* 8.
- Delattre, Pierre. 1966a. « A comparison of syllable length conditioning among languages ». *IRAL - International Review of Applied Linguistics in Language Teaching* 4(1-4). doi: [10.1515/iral.1966.4.1-4.183](https://doi.org/10.1515/iral.1966.4.1-4.183).
- Delattre, Pierre. 1966b. « Les dix intonations de base du français ». *French review* 1-14.
- Dellwo, Volker. 2006. « Rhythm and Speech Rate: A Variation Coefficient for deltaC ». doi: [10.5167/UZH-111789](https://doi.org/10.5167/UZH-111789).
- Dewey, John A., et Günther Knoblich. 2014. « Do Implicit and Explicit Measures of the Sense of Agency Measure the Same Thing? » édité par S. Kotz. *PLoS ONE* 9(10):e110118. doi: [10.1371/journal.pone.0110118](https://doi.org/10.1371/journal.pone.0110118).
- Di Cristo, Albert. 1978. « De la microprosodie à l'intonosyntaxe ». PhD Thesis.
- Di Cristo, Albert. 2003. « De la métrique et du rythme de la parole ordinaire: l'exemple du français ». *Semen* (16). doi: [10.4000/semen.2944](https://doi.org/10.4000/semen.2944).
- Di Cristo, Albert. 2011. « Une Approche Intégrative Des Relations de l'accentuation Au Phrasé Prosodique Du Français ». *Journal of French Language Studies* 21(1):73-95. doi: [10.1017/S0959269510000505](https://doi.org/10.1017/S0959269510000505).
- Di Cristo, Albert. 2016. *Les musiques du français parlé: essais sur l'accentuation, la métrique, le rythme, le phrasé prosodique et l'intonation du français contemporain*. Berlin: De Gruyter.
- Di Cristo, Albert, et Daniel Hirst. 1993. « Rythme syllabique, rythme mélodique et représentation hiérarchique de la prosodie du français ». *Travaux de l'Institut de Phonétique d'Aix* 15:9-24.
- Di Cristo, Albert, et Daniel Hirst. 1996. « Vers une typologie des unités intonatives du français ». *XXIème JEP, Avignon* 219-22.
- Doelling, Keith B., Luc H. Arnal, Oded Ghitza, et David Poeppel. 2014. « Acoustic Landmarks Drive Delta-Theta Oscillations to Enable Speech Comprehension by Facilitating Perceptual Parsing ». *NeuroImage* 85:761-68. doi: [10.1016/j.neuroimage.2013.06.035](https://doi.org/10.1016/j.neuroimage.2013.06.035).
- Fairbanks, Grant. 1955. « Selective Vocal Effects Of Delayed Auditory Feedback ». *Journal of Speech and Hearing Disorders* 20(4):333-46. doi: [10.1044/jshd.2004.333](https://doi.org/10.1044/jshd.2004.333).
- Falk, Simone, et Christopher T. Kello. 2017. « Hierarchical Organization in the Temporal Structure of Infant-Direct Speech and Song ». *Cognition* 163:80-86. doi: [10.1016/j.cognition.2017.02.017](https://doi.org/10.1016/j.cognition.2017.02.017).
- Farrer, C., M. Bouchereau, M. Jeannerod, et N. Franck. 2008. « Effect of Distorted Visual Feedback on the Sense of Agency ». *Behavioural Neurology* 19(1-2):53-57. doi: [10.1155/2008/425267](https://doi.org/10.1155/2008/425267).
- Fónagy, Ivan, et Judith Fónagy. 1983. « L'intonation et l'organisation du discours ». *Bulletin de la Société de Linguistique de Paris* 78(1):161-209.

- Fougeron, Cécile. 2005. « La phonologie articulatoire: une introduction ». *N. Nguyen, S. Wauquier-Gravelines et J. Durand, éditeurs: Phonologie et phonétique: Forme et substance, Cognition et Traitement de l'Information* 265-90.
- Fraisse, Paul. 1956. *Les structures rythmiques: étude psychologique*. Publications universitaires de Louvain.
- Franken, Matthias K., Daniel J. Acheson, James M. McQueen, Peter Hagoort, et Frank Eisner. 2018. « Opposing and Following Responses in Sensorimotor Speech Control: Why Responses Go Both Ways ». *Psychonomic Bulletin & Review* 25(4):1458-67. doi: [10.3758/s13423-018-1494-x](https://doi.org/10.3758/s13423-018-1494-x).
- Franken, Matthias K., Daniel J. Acheson, James M. McQueen, Peter Hagoort, et Frank Eisner. 2019. « Consistency Influences Altered Auditory Feedback Processing ». *Quarterly Journal of Experimental Psychology* 72(10):2371-79. doi: [10.1177/1747021819838939](https://doi.org/10.1177/1747021819838939).
- Franken, Matthias K., Robert J. Hartsuiker, Petter Johansson, Lars Hall, et Andreas Lind. 2021. « Speaking with an Alien Voice: Flexible Sense of Agency during Vocal Production. » *Journal of Experimental Psychology: Human Perception and Performance* 47(4):479-94. doi: [10.1037/xhp0000799](https://doi.org/10.1037/xhp0000799).
- Franken, Matthias K., Robert J. Hartsuiker, Petter Johansson, Lars Hall, et Andreas Lind. 2022. « Drifting Pitch Awareness after Exposure to Altered Auditory Feedback ». *Attention, Perception, & Psychophysics* 84(6):2027-39. doi: [10.3758/s13414-022-02441-2](https://doi.org/10.3758/s13414-022-02441-2).
- Gentilucci, Maurizio, et Paolo Bernardis. 2007. « Imitation during Phoneme Production ». *Neuropsychologia* 45(3):608-15. doi: [10.1016/j.neuropsychologia.2006.04.004](https://doi.org/10.1016/j.neuropsychologia.2006.04.004).
- Ghitza, Oded, et Steven Greenberg. 2009. « On the Possible Role of Brain Rhythms in Speech Perception: Intelligibility of Time-Compressed Speech with Periodic and Aperiodic Insertions of Silence ». *Phonetica* 66(1-2):113-26. doi: [10.1159/000208934](https://doi.org/10.1159/000208934).
- Giraud, Anne-Lise, et David Poeppel. 2012. « Cortical Oscillations and Speech Processing: Emerging Computational Principles and Operations ». *Nature Neuroscience* 15(4):511-17. doi: [10.1038/nn.3063](https://doi.org/10.1038/nn.3063).
- Goldstein, Louis. 2019. « The Role of Temporal Modulation in Sensorimotor Interaction ». *Frontiers in Psychology* 10:2608. doi: [10.3389/fpsyg.2019.02608](https://doi.org/10.3389/fpsyg.2019.02608).
- Goldstein, Louis, et Carol A. Fowler. 2003. « Articulatory phonology: A phonology for public language use ». *Phonetics and phonology in language comprehension and production: Differences and similarities* 159-207. doi: <https://doi.org/10.1515/9783110895094>.
- Grabe, Esther, et Ee Ling Low. 2002. « Durational variability in speech and the Rhythm Class Hypothesis ». P. 515-46 in *Laboratory Phonology 7*, édité par C. Gussenhoven et N. Warner. Mouton de Gruyter.
- Grabe, Esther, Brechtje Post, et Ian Watson. 1999. « The acquisition of rhythmic patterns in English and French ». P. 1201-4 in *Proceedings of the 14th International Congress of Phonetic Sciences*. University of California Berkeley, CA.

- Guenther, Frank H., Satrajit S. Ghosh, et Jason A. Tourville. 2006. « Neural modeling and imaging of the cortical interactions underlying syllable production ». *Brain and Language* 96(3):280-301. doi: [10.1016/j.bandl.2005.06.001](https://doi.org/10.1016/j.bandl.2005.06.001).
- Haggard, Patrick. 2017. « Sense of Agency in the Human Brain ». *Nature Reviews Neuroscience* 18(4):196-207. doi: [10.1038/nrn.2017.14](https://doi.org/10.1038/nrn.2017.14).
- Haggard, Patrick, et Valerian Chambon. 2012. « Sense of Agency ». *Current Biology* 22(10):R390-92. doi: [10.1016/j.cub.2012.02.040](https://doi.org/10.1016/j.cub.2012.02.040).
- Haggard, Patrick, Sam Clark, et Jeri Kalogeras. 2002. « Voluntary Action and Conscious Awareness ». *Nature Neuroscience* 5(4):382-85. doi: [10.1038/nm827](https://doi.org/10.1038/nm827).
- Hain, Timothy C., Theresa A. Burnett, Swathi Kiran, Charles R. Larson, Shajila Singh, et Mary K. Kenney. 2000. « Instructing subjects to make a voluntary response reveals the presence of two components to the audio-vocal reflex ». *Experimental Brain Research* 130(2):133-41. doi: [10.1007/s002219900237](https://doi.org/10.1007/s002219900237).
- Halle, Morris, et Jean-Roger Vergnaud. 1987. « Stress and the cycle ». *Linguistic inquiry* 18(1):45-84.
- Hallé, Pierre A., Bénédicte De Boysson-Bardies, et Marilyn M. Vihman. 1991. « Beginnings of Prosodic Organization: Intonation and Duration Patterns of Disyllables Produced by Japanese and French Infants ». *Language and Speech* 34(4):299-318. doi: [10.1177/002383099103400401](https://doi.org/10.1177/002383099103400401).
- Hayes, Bruce. 1984. « The phonology of rhythm in English ». *Linguistic Inquiry* 15(1):33-74.
- Hayes, Bruce Philip. 1980. « A metrical theory of stress rules ». PhD Thesis, Massachusetts Institute of Technology.
- Hermes, Anne, Sam Tilsen, et Rachid Ridouane. 2020. « Cross-linguistic timing contrast in geminates: A rate-independent perspective ». P. 52-55 in *12th International Seminar on Speech Production (ISSP2020)*.
- Hogg, Richard M., et Christopher Bayston McCully. 1987. *Metrical phonology: a course book*. Cambridge University Press.
- Houde, John F., et Michael I. Jordan. 1998. « Sensorimotor Adaptation in Speech Production ». *Science* 279(5354):1213-16. doi: [10.1126/science.279.5354.1213](https://doi.org/10.1126/science.279.5354.1213).
- Houde, John F., et Michael I. Jordan. 2002. « Sensorimotor Adaptation of Speech I: Compensation and Adaptation ». *Journal of Speech, Language, and Hearing Research* 45(2):295-310. doi: [10.1044/1092-4388\(2002/023\)](https://doi.org/10.1044/1092-4388(2002/023)).
- Houde, John F., et Srikantan S. Nagarajan. 2011. « Speech Production as State Feedback Control ». *Frontiers in Human Neuroscience* 5. doi: [10.3389/fnhum.2011.00082](https://doi.org/10.3389/fnhum.2011.00082).
- Houde, John F., Srikantan S. Nagarajan, Kensuke Sekihara, et Michael M. Merzenich. 2002. « Modulation of the Auditory Cortex during Speech: An MEG Study ». *Journal of Cognitive Neuroscience* 14(8):1125-38. doi: [10.1162/089892902760807140](https://doi.org/10.1162/089892902760807140).

- Howell, Peter. 1990. « Changes in Voice Level Caused by Several Forms of Altered Feedback in Fluent Speakers and Stutterers ». *Language and Speech* 33(4):325-38. doi: [10.1177/002383099003300402](https://doi.org/10.1177/002383099003300402).
- Howell, Peter. 2004. « Effects of delayed auditory feedback and frequency-shifted feedback on speech control and some potentials for future development of prosthetic aids for stammering ». *Stammering research: an on-line journal published by the British Stammering Association* 1(1):31.
- Howell, Peter, et Alexandra Archer. 1984. « Susceptibility to the Effects of Delayed Auditory Feedback ». *Perception & Psychophysics* 36(3):296-302. doi: [10.3758/BF03206371](https://doi.org/10.3758/BF03206371).
- Huang, Norden E., Zhaohua Wu, Steven R. Long, Kenneth C. Arnold, Xianyao Chen, et Karin Blank. 2009. « On Instantaneous Frequency ». *Advances in Adaptive Data Analysis* 01(02):177-229. doi: [10.1142/S1793536909000096](https://doi.org/10.1142/S1793536909000096).
- Jankowski, Ludovic, Corine Astésano, et Albert Di Cristo. 1999. « The initial rhythmic accent in French: Acoustic data and perceptual investigation ». P. 257-60 in *Proceedings of the 14th International Congress of Phonetic Sciences*. Vol. 1.
- Jasmin, Kyle M., Carolyn McGettigan, Zarinah K. Agnew, Nadine Lavan, Oliver Josephs, Fred Cummins, et Sophie K. Scott. 2016. « Cohesion and Joint Speech: Right Hemisphere Contributions to Synchronized Vocal Production ». *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 36(17):4669-80. doi: [10.1523/JNEUROSCI.4075-15.2016](https://doi.org/10.1523/JNEUROSCI.4075-15.2016).
- Jones, Jeffery A., et K. G. Munhall. 2003. « Learning to Produce Speech with an Altered Vocal Tract: The Role of Auditory Feedback ». *The Journal of the Acoustical Society of America* 113(1):532-43. doi: [10.1121/1.1529670](https://doi.org/10.1121/1.1529670).
- Jun, Sun-Ah. 2012. « Prosodic typology revisited: Adding macro-rhythm ». in *Speech Prosody 2012*.
- Jun, Sun-Ah, et Cécile Fougeron. 2000. « A phonological model of French intonation ». P. 209-42 in *Intonation*. Vol. 15, *Text, Speech and Language Technology*, édité par A. Botinis. Dordrecht: Springer Netherlands.
- Kalveram, Karl Theodor, et Lutz Jäncke. 1989. « Vowel Duration and Voice Onset Time for Stressed and Nonstressed Syllables in Stutterers under Delayed Auditory Feedback Condition ». *Folia Phoniatrica et Logopaedica* 41(1):30-42. doi: [10.1159/000265930](https://doi.org/10.1159/000265930).
- Karlin, Robin, Chris Naber, et Benjamin Parrell. 2021. « Auditory Feedback Is Used for Adaptation and Compensation in Speech Timing ». *Journal of Speech, Language, and Hearing Research* 64(9):3361-81. doi: [10.1044/2021\\_JSLHR-21-00021](https://doi.org/10.1044/2021_JSLHR-21-00021).
- Kelso, J. Scott, Betty Tuller, E. Vatikiotis-Bateson, et Carol A. Fowler. 1984. « Functionally Specific Articulatory Cooperation Following Jaw Perturbations during Speech: Evidence for Coordinative Structures. » *Journal of Experimental Psychology: Human Perception and Performance* 10(6):812-32. doi: [10.1037/0096-1523.10.6.812](https://doi.org/10.1037/0096-1523.10.6.812).
- Köhler, Wolfgang. 1967. « Gestalt Psychology ». *Psychologische Forschung* 31(1):XVIII-XXX. doi: [10.1007/BF00422382](https://doi.org/10.1007/BF00422382).

- Kröger, Bernd J., Jim Kannampuzha, et Christiane Neuschaefer-Rube. 2009. « Towards a Neurocomputational Model of Speech Production and Perception ». *Speech Communication* 51(9):793-809. doi: [10.1016/j.specom.2008.08.002](https://doi.org/10.1016/j.specom.2008.08.002).
- Kruckenberger, Anita, et Gunar Fant. 1995. « Notes on syllable duration in French and Swedish ». P. 158-61 in *Proc. ICPbS*. Vol. 95.
- Lametti, D. R., S. M. Nasir, et D. J. Ostry. 2012. « Sensory Preference in Speech Production Revealed by Simultaneous Alteration of Auditory and Somatosensory Feedback ». *Journal of Neuroscience* 32(27):9351-58. doi: [10.1523/JNEUROSCI.0404-12.2012](https://doi.org/10.1523/JNEUROSCI.0404-12.2012).
- Lancia, Leonardo, Georgy Krasovitsky, et Franziska Stuntebeck. 2019. « Coordinative Patterns Underlying Cross-Linguistic Rhythmic Differences ». *Journal of Phonetics* 72:66-80. doi: [10.1016/j.wocn.2018.08.004](https://doi.org/10.1016/j.wocn.2018.08.004).
- Lancia, Leonardo, Jinyu Li, et Cécile Fougeron. 2023. « How speech rate, syllabic complexity and diversity affect the emergence of speech rhythm in speeded syllable repetition ». in *20th International Congress of Phonetic Sciences (ICPhS)*. Prague, Czech Republic.
- Lancia, Leonardo, et Benjamin Rosenbaum. 2018. « Coupling Relations Underlying the Production of Speech Articulator Movements and Their Invariance to Speech Rate ». *Biological Cybernetics* 112(3):253-76. doi: [10.1007/s00422-018-0749-y](https://doi.org/10.1007/s00422-018-0749-y).
- Larson, Charles R., Kenneth W. Altman, Hanjun Liu, et Timothy C. Hain. 2008. « Interactions between Auditory and Somatosensory Feedback for Voice F0 Control ». *Experimental Brain Research* 187(4):613-21. doi: [10.1007/s00221-008-1330-z](https://doi.org/10.1007/s00221-008-1330-z).
- Larson, Charles R., Jean Sun, et Timothy C. Hain. 2007. « Effects of Simultaneous Perturbations of Voice Pitch and Loudness Feedback on Voice F0 and Amplitude Control ». *The Journal of the Acoustical Society of America* 121(5):2862-72. doi: [10.1121/1.2715657](https://doi.org/10.1121/1.2715657).
- Leong, Victoria, et Usha Goswami. 2014. « Assessment of Rhythmic Entrainment at Multiple Timescales in Dyslexia: Evidence for Disruption to Syllable Timing ». *Hearing Research* 308:141-61. doi: [10.1016/j.heares.2013.07.015](https://doi.org/10.1016/j.heares.2013.07.015).
- Leong, Victoria, et Usha Goswami. 2015. « Acoustic-Emergent Phonology in the Amplitude Envelope of Child-Directed Speech » édité par F. Dick. *PLOS ONE* 10(12):e0144411. doi: [10.1371/journal.pone.0144411](https://doi.org/10.1371/journal.pone.0144411).
- Leong, Victoria, Marina Kalashnikova, Denis Burnham, et Usha Goswami. 2014. « Infant-directed speech enhances temporal rhythmic structure in the envelope ». in *Fifteenth Annual Conference of the International Speech Communication Association*.
- Lieberman, Mark, et Alan Prince. 1977. « On stress and linguistic rhythm ». *Linguistic inquiry* 8(2):249-336.
- Lin, Yuhan, Yao Yao, et Jin Luo. 2021. « Phonetic Accommodation of Tone: Reversing a Tone Merger-in-Progress via Imitation ». *Journal of Phonetics* 87:101060. doi: [10.1016/j.wocn.2021.101060](https://doi.org/10.1016/j.wocn.2021.101060).

- Liu, Hanjun, et Charles R. Larson. 2007. « Effects of Perturbation Magnitude and Voice F Level on the Pitch-Shift Reflex ». *The Journal of the Acoustical Society of America* 122(6):3671-77. doi: [10.1121/1.2800254](https://doi.org/10.1121/1.2800254).
- Liu, Hanjun, Emily Q. Wang, Leo Verhagen Metman, et Charles R. Larson. 2012. « Vocal Responses to Perturbations in Voice Auditory Feedback in Individuals with Parkinson's Disease » édité par J. Yan. *PLoS ONE* 7(3):e33629. doi: [10.1371/journal.pone.0033629](https://doi.org/10.1371/journal.pone.0033629).
- Low, Ee Ling, Esther Grabe, et Francis Nolan. 2000. « Quantitative Characterizations of Speech Rhythm: Syllable-Timing in Singapore English ». *Language and Speech* 43(4):377-401. doi: [10.1177/00238309000430040301](https://doi.org/10.1177/00238309000430040301).
- MacNeilage, Peter F. 1998. « The Frame/Content Theory of Evolution of Speech Production ». *Behavioral and Brain Sciences* 21(4):499-511. doi: [10.1017/S0140525X98001265](https://doi.org/10.1017/S0140525X98001265).
- Maddieson, Ian. 1985. « Phonetic cues to syllabification ». *Phonetic linguistics: Essays in honor of Peter Ladefoged* 203:221.
- Mermelstein, Paul. 1975. « Automatic Segmentation of Speech into Syllabic Units ». *The Journal of the Acoustical Society of America* 58(4):880-83. doi: [10.1121/1.380738](https://doi.org/10.1121/1.380738).
- Meynadier, Yohann, Michel Pitermann, et Alain Marchal. 1998. « Effects of Contrastive Focal Accent on Linguopalatal Articulation and Coarticulation in the French [Kskl] Cluster. » doi: [10.13140/RG.2.1.3595.6720](https://doi.org/10.13140/RG.2.1.3595.6720).
- Miall, R. C., et D. M. Wolpert. 1996. « Forward Models for Physiological Motor Control ». *Neural Networks* 9(8):1265-79. doi: [10.1016/S0893-6080\(96\)00035-4](https://doi.org/10.1016/S0893-6080(96)00035-4).
- Michelas, Amandine, et Sophie Dufour. 2019. « Are Prosodic Variants Stored in the French Mental Lexicon? » *Experimental Psychology* 66(6):393-401. doi: [10.1027/1618-3169/a000462](https://doi.org/10.1027/1618-3169/a000462).
- Michelas, Amandine, Ulrich Hans Frauenfelder, Daniele Schön, et Sophie Dufour. 2016. « How Deaf Are French Speakers to Stress? » *The Journal of the Acoustical Society of America* 139(3):1333-42. doi: [10.1121/1.4944574](https://doi.org/10.1121/1.4944574).
- Michelas, Amandine, et Noël Nguyen. 2011. « Uncovering the effect of imitation on tonal patterns of French Accentual Phrases ». P. 973-76 in *Interspeech 12*. Florence, Italy.
- Mitsuya, Takashi, Fabienne Samson, Lucie Ménard, et Kevin G. Munhall. 2013. « Language Dependent Vowel Representation in Speech Production ». *The Journal of the Acoustical Society of America* 133(5):2993-3003. doi: [10.1121/1.4795786](https://doi.org/10.1121/1.4795786).
- Mukherjee, Sankar, Thierry Legou, Leonardo Lancia, Pauline Hilt, Alice Tomassini, Luciano Fadiga, Alessandro d'Ausilio, Leonardo Badino, et Noël Nguyen. 2018. « Analyzing vocal tract movements during speech accommodation ». P. Paper-2084 in *Interspeech 2018*. ISCA.
- Neuhoff, John G., Rebecca Knight, et Joseph Wayand. 2002. « Pitch change, sonification, and musical expertise: Which way is up? » in *International Conference on Auditory Display (ICAD)*. Georgia Institute of Technology.
- Niziolek, C. A., et F. H. Guenther. 2013. « Vowel Category Boundaries Enhance Cortical and Behavioral Responses to Speech Feedback Alterations ». *Journal of Neuroscience* 33(29):12090-98. doi: [10.1523/JNEUROSCI.1008-13.2013](https://doi.org/10.1523/JNEUROSCI.1008-13.2013).

- Ohata, Ryu, Tomohisa Asai, Shu Imaizumi, et Hiroshi Imamizu. 2022. « I Hear My Voice; Therefore I Spoke: The Sense of Agency Over Speech Is Enhanced by Hearing One's Own Voice ». *Psychological Science* 33(8):1226-39. doi: [10.1177/09567976211068880](https://doi.org/10.1177/09567976211068880).
- Oschkinat, Miriam, et Philip Hoole. 2022. « Reactive Feedback Control and Adaptation to Perturbed Speech Timing in Stressed and Unstressed Syllables ». *Journal of Phonetics* 91:101133. doi: [10.1016/j.wocn.2022.101133](https://doi.org/10.1016/j.wocn.2022.101133).
- Pardo, Jennifer S. 2006. « On Phonetic Convergence during Conversational Interaction ». *The Journal of the Acoustical Society of America* 119(4):2382-93. doi: [10.1121/1.2178720](https://doi.org/10.1121/1.2178720).
- Pardo, Jennifer S., Rachel Gibbons, Alexandra Suppes, et Robert M. Krauss. 2012. « Phonetic Convergence in College Roommates ». *Journal of Phonetics* 40(1):190-97. doi: [10.1016/j.wocn.2011.10.001](https://doi.org/10.1016/j.wocn.2011.10.001).
- Pardo, Jennifer S., Kelly Jordan, Rolliene Mallari, Caitlin Scanlon, et Eva Lewandowski. 2013. « Phonetic Convergence in Shadowed Speech: The Relation between Acoustic and Perceptual Measures ». *Journal of Memory and Language* 69(3):183-95. doi: [10.1016/j.jml.2013.06.002](https://doi.org/10.1016/j.jml.2013.06.002).
- Parrell, Benjamin, Adam C. Lammert, Gregory Ciccarelli, et Thomas F. Quatieri. 2019. « Current Models of Speech Motor Control: A Control-Theoretic Overview of Architectures and Properties ». *The Journal of the Acoustical Society of America* 145(3):1456-81. doi: [10.1121/1.5092807](https://doi.org/10.1121/1.5092807).
- Parrell, Benjamin, Vikram Ramanarayanan, Srikantan Nagarajan, et John Houde. 2019. « The FACTS Model of Speech Motor Control: Fusing State Estimation and Task-Based Control » édité par A. M. Haith. *PLOS Computational Biology* 15(9):e1007321. doi: [10.1371/journal.pcbi.1007321](https://doi.org/10.1371/journal.pcbi.1007321).
- Patri, Jean-François, Julien Diard, et Pascal Perrier. 2015. « Optimal Speech Motor Control and Token-to-Token Variability: A Bayesian Modeling Approach ». *Biological Cybernetics* 109(6):611-26. doi: [10.1007/s00422-015-0664-4](https://doi.org/10.1007/s00422-015-0664-4).
- Payne, Elinor, Brechtje Post, Lluïsa Astruc, Pilar Prieto, et Maria del Mar Vanrell. 2012. « Measuring Child Rhythm ». *Language and Speech* 55(2):203-29. doi: [10.1177/0023830911417687](https://doi.org/10.1177/0023830911417687).
- Perrier, Pascal, Liang Ma, et Yohan Payan. 2006. « Modeling the production of VCV sequences via the inversion of a biomechanical model of the tongue ». doi: [10.48550/ARXIV.PHYSICS/0610170](https://doi.org/10.48550/ARXIV.PHYSICS/0610170).
- Pickering, Martin J., et Simon Garrod. 2004. « Toward a Mechanistic Psychology of Dialogue ». *Behavioral and Brain Sciences* 27(02). doi: [10.1017/S0140525X04000056](https://doi.org/10.1017/S0140525X04000056).
- Pickering, Martin J., et Simon Garrod. 2013. « An Integrated Theory of Language Production and Comprehension ». *Behavioral and Brain Sciences* 36(4):329-47. doi: [10.1017/S0140525X12001495](https://doi.org/10.1017/S0140525X12001495).
- Pike, Kenneth L. 1945. *The Intonation of American English*. Ann Arbor: University of Michigan Press.

- Poeppl, David, et M. Florencia Assaneo. 2020. « Speech Rhythms and Their Neural Foundations ». *Nature Reviews Neuroscience* 21(6):322-34. doi: [10.1038/s41583-020-0304-4](https://doi.org/10.1038/s41583-020-0304-4).
- Polyanskaya, Leona, Maria Grazia Busà, et Mikhail Ordin. 2020. « Capturing Cross-Linguistic Differences in Macro-Rhythm: The Case of Italian and English ». *Language and Speech* 63(2):242-63. doi: [10.1177/0023830919835849](https://doi.org/10.1177/0023830919835849).
- Prince, Alan S. 1983. « Relating to the Grid ». *Linguistic Inquiry* 14(1):19-100.
- Ramus, Franck, Marina Nespore, et Jacques Mehler. 1999. « Correlates of Linguistic Rhythm in the Speech Signal ». *Cognition* 73(3):265-92. doi: [10.1016/S0010-0277\(99\)00058-X](https://doi.org/10.1016/S0010-0277(99)00058-X).
- Rietveld, ACM, et Ulrich Hans Frauenfelder. 1987. « The effect of syllable structure on vowel duration ». in *11th International Congress of Phonetic Sciences*. Tallinn, Estonia.
- Robinson, George M. 1972. « The Delayed Auditory Feedback Effect Is a Function of Speech Rate ». *Journal of Experimental Psychology* 95(1):1-5. doi: [10.1037/h0033268](https://doi.org/10.1037/h0033268).
- Rochet-Capellan, Amélie, et Jean-Luc Schwartz. 2007. « An Articulatory Basis for the Labial-to-Coronal Effect: /Pata/ Seems a More Stable Articulatory Pattern than /Tapa/ ». *The Journal of the Acoustical Society of America* 121(6):3740. doi: [10.1121/1.2734497](https://doi.org/10.1121/1.2734497).
- Rosenblum, M., A. Pikovsky, J. Kurths, C. Schäfer, et P. A. Tass. 2001. « Phase Synchronization: From Theory to Data Analysis ». P. 279-321 in *Handbook of Biological Physics*. Vol. 4. Elsevier.
- Rossi, M., et A. Di Cristo. 1980. « Un modèle de détection automatique des frontières intonatives et syntaxiques ». *Journées d'Etude sur la Parole* 213-38.
- Rossi, Mario. 1980. « Le français, langue sans accent? » *Studia Phonetica Montréal* 15:13-51.
- Rossi, Mario. 1985. « L'intonation et l'organisation de l'énoncé ». *Phonetica* 42(2-3):135-53. doi: [10.1159/000261744](https://doi.org/10.1159/000261744).
- Sakoe, H., et S. Chiba. 1978. « Dynamic Programming Algorithm Optimization for Spoken Word Recognition ». *IEEE Transactions on Acoustics, Speech, and Signal Processing* 26(1):43-49. doi: [10.1109/TASSP.1978.1163055](https://doi.org/10.1109/TASSP.1978.1163055).
- Saltzman, Elliot, et J. A. Kelso. 1987. « Skilled Actions: A Task-Dynamic Approach ». *Psychological Review* 94(1):84-106. doi: [10.1037/0033-295X.94.1.84](https://doi.org/10.1037/0033-295X.94.1.84).
- Saltzman, Elliot L., et Kevin G. Munhall. 1989. « A Dynamical Approach to Gestural Patterning in Speech Production ». *Ecological Psychology* 1(4):333-82. doi: [10.1207/s15326969eco0104\\_2](https://doi.org/10.1207/s15326969eco0104_2).
- Sato, Marc, Krystyna Grabski, Maëva Garnier, Lionel Granjon, Jean-Luc Schwartz, et Noël Nguyen. 2013. « Converging toward a common speech code: imitative and perceptuo-motor recalibration processes in speech production ». *Frontiers in Psychology* 4. doi: [10.3389/fpsyg.2013.00422](https://doi.org/10.3389/fpsyg.2013.00422).
- Shiller, Douglas M., Marc Sato, Vincent L. Gracco, et Shari R. Baum. 2009. « Perceptual Recalibration of Speech Sounds Following Speech Motor Learning ». *The Journal of the Acoustical Society of America* 125(2):1103-13. doi: [10.1121/1.3058638](https://doi.org/10.1121/1.3058638).

- Shiraishi, Masahiro, et Sotaro Shimada. 2021. « Inter-Brain Synchronization during a Cooperative Task Reflects the Sense of Joint Agency ». *Neuropsychologia* 154:107770. doi: [10.1016/j.neuropsychologia.2021.107770](https://doi.org/10.1016/j.neuropsychologia.2021.107770).
- Smith, Caroline L., Donna Erickson, et Christophe Savariaux. 2019. « Articulatory and Acoustic Correlates of Prominence in French: Comparing L1 and L2 Speakers ». *Journal of Phonetics* 77:100938. doi: [10.1016/j.wocn.2019.100938](https://doi.org/10.1016/j.wocn.2019.100938).
- Stetson, R. H. 1951. *Motor phonetics: a study of speech movements in action*. Amsterdam: North-Holland Publishing Co.
- Stuart, Andrew, et Joseph Kalinowski. 2015. « Effect of Delayed Auditory Feedback, Speech Rate, and Sex on Speech Production ». *Perceptual and Motor Skills* 120(3):747-65. doi: [10.2466/23.25.PMS.120v17x2](https://doi.org/10.2466/23.25.PMS.120v17x2).
- Tilsen, Sam, et Amalia Arvaniti. 2013. « Speech Rhythm Analysis with Decomposition of the Amplitude Envelope: Characterizing Rhythmic Patterns within and across Languages ». *The Journal of the Acoustical Society of America* 134(1):628-39. doi: [10.1121/1.4807565](https://doi.org/10.1121/1.4807565).
- Tourville, Jason A., et Frank H. Guenther. 2011. « The DIVA Model: A Neural Theory of Speech Acquisition and Production ». *Language and Cognitive Processes* 26(7):952-81. doi: [10.1080/01690960903498424](https://doi.org/10.1080/01690960903498424).
- Troubetzkoy, Nikolai. 1949. « Principes de Phonologie (traduit par J. Cantineau). Paris, Klincksieck ». *Revue des Études Grecques* 62(291):463-66.
- Tuller, Betty, et J. A. S. Kelso. 1991. « The Production and Perception of Syllable Structure ». *Journal of Speech, Language, and Hearing Research* 34(3):501-8. doi: [10.1044/jshr.3403.501](https://doi.org/10.1044/jshr.3403.501).
- Venables, W. N., et B. D. Ripley. 2002. *Modern Applied Statistics with S*. New York, NY: Springer New York.
- Wioland, François. 1984. « Organisation temporelle des structures rythmiques du français parlé ». *Bulletin de la Section de Linguistique de la Faculté des Lettres de Lausanne* (6):293-322.
- Wolpert, Daniel M., Zoubin Ghahramani, et Michael I. Jordan. 1995. « An Internal Model for Sensorimotor Integration ». *Science* 269(5232):1880-82. doi: [10.1126/science.7569931](https://doi.org/10.1126/science.7569931).
- Yates, Aubrey J. 1963. « Delayed Auditory Feedback ». *Psychological Bulletin* 60(3):213-32. doi: [10.1037/h0044155](https://doi.org/10.1037/h0044155).
- Zheng, Zane Z., Ewen N. MacDonald, Kevin G. Munhall, et Ingrid S. Johnsrude. 2011. « Perceiving a Stranger's Voice as Being One's Own: A 'Rubber Voice' Illusion? » édité par A. Sirigu. *PLoS ONE* 6(4):e18655. doi: [10.1371/journal.pone.0018655](https://doi.org/10.1371/journal.pone.0018655).
- Zimmermann, G., C. Brown, J. A. S. Kelso, R. Hurtig, et K. Forrest. 1988. « The Association between Acoustic and Articulatory Events in a Delayed Auditory Feedback Paradigm ». *Journal of Phonetics* 16(4):437-51. doi: [10.1016/S0095-4470\(19\)30520-0](https://doi.org/10.1016/S0095-4470(19)30520-0).

# ANNEXE

Tableau 1 Résumé des résultats du modèle linéaire mixte portant sur l'effet du DAF sur la durée des voyelles dans différentes positions de la **phrase 1** dans l'expérience I (cf. chapitre IV 1.1)

**Modèle** :  $Durée \sim DAF + Voyelle + Position + DAF:Voyelle + DAF:Position + Voyelle:Position + (1 | loc.:DAF) + (1 | loc.:Voyelle) + (1 | loc.:Position)^{14}$

**Référence** : la durée moyenne des voyelles dans répétitions de phrases en Position 1 dans les essais sans DAF

	Estimée	Écart-type	DL <sup>15</sup>	t-valeur	p-valeur
<b>Intercept</b>	0,098	0,006	60,870	16,200	<b>&lt;0,001</b>
V2	0,036	0,013	37,230	2,829	<b>0,007</b>
V3	-0,049	0,013	37,230	-3,845	<b>&lt;0,001</b>
V7	0,040	0,013	37,230	3,159	<b>0,003</b>
V5:Position2	0,006	0,003	3266,000	2,361	<b>0,018</b>
V2:Position3	0,006	0,003	3266,000	2,513	<b>0,012</b>
V5:Position3	0,008	0,003	3266,000	3,225	<b>0,001</b>
V7:Position3	-0,021	0,003	3266,000	-8,389	<b>&lt;0,001</b>
DAF60	0,017	0,004	51,150	4,254	<b>&lt;0,001</b>
DAF90	0,017	0,004	48,310	4,377	<b>&lt;0,001</b>
DAF120	0,020	0,004	44,710	5,135	<b>&lt;0,001</b>
DAF150	0,025	0,004	47,790	6,247	<b>&lt;0,001</b>
DAF180	0,024	0,004	49,690	5,982	<b>&lt;0,001</b>
DAF120:Position2	0,014	0,004	3299,000	3,984	<b>&lt;0,001</b>
DAF180:Position2	0,011	0,004	3299,000	2,739	<b>0,006</b>
DAF60:V2	0,009	0,004	3265,000	2,505	<b>0,012</b>
DAF90:V2	0,007	0,004	3265,000	2,101	<b>0,036</b>
DAF120:V2	0,008	0,004	3265,000	2,327	<b>0,020</b>
DAF150:V2	0,008	0,004	3266,000	2,121	<b>0,034</b>
DAF60:V3	-0,011	0,004	3265,000	-3,049	<b>0,002</b>
DAF90:V3	-0,015	0,004	3265,000	-4,338	<b>&lt;0,001</b>
DAF120:V3	-0,014	0,004	3265,000	-3,969	<b>&lt;0,001</b>
DAF150:V3	-0,019	0,004	3266,000	-5,400	<b>&lt;0,001</b>
DAF180:V3	-0,020	0,004	3266,000	-5,562	<b>&lt;0,001</b>
DAF60:V4	-0,007	0,004	3265,000	-2,054	<b>0,040</b>
DAF90:V4	-0,010	0,004	3265,000	-2,859	<b>0,004</b>
DAF120:V4	-0,009	0,004	3265,000	-2,602	<b>0,009</b>
DAF150:V4	-0,011	0,004	3266,000	-3,128	<b>0,002</b>
DAF180:V4	-0,012	0,004	3266,000	-3,378	<b>0,001</b>
DAF60:V5	0,008	0,004	3265,000	2,180	<b>0,029</b>
DAF90:V5	0,013	0,004	3265,000	3,764	<b>&lt;0,001</b>
DAF120:V5	0,015	0,004	3265,000	4,166	<b>&lt;0,001</b>

<sup>14</sup> La terme « Position » est utilisé de manière cohérente dans tous les tableaux présentés dans cette thèse pour représenter « la position de la phrase dans un essai expérimental ». L'abréviation « loc. » est utilisée de manière cohérente dans tous les tableaux présentés dans cette thèse et fait référence à la variable « locutrice ».

<sup>15</sup> Dans ce tableau, « DL » est l'acronyme utilisé pour représenter le « degré de liberté ». Cette abréviation est employée de manière cohérente dans l'ensemble des tableaux présentés dans cette thèse.

	Estimée	Écart-type	DL <sup>15</sup>	t-valeur	p-valeur
DAF150:V5	0,015	0,004	3266,000	4,253	<0,001
DAF180:V5	0,022	0,004	3266,000	6,055	<0,001
DAF180:V6	-0,008	0,004	3266,000	-2,213	0,027
DAF60:V7	0,018	0,004	3265,000	5,075	<0,001
DAF90:V7	0,025	0,004	3265,000	7,098	<0,001
DAF120:V7	0,027	0,004	3265,000	7,613	<0,001
DAF150:V7	0,032	0,004	3266,000	9,014	<0,001
DAF180:V7	0,038	0,004	3266,000	10,627	<0,001

Tableau 2 Résumé des résultats du modèle linéaire mixte portant sur l'effet du DAF sur la durée des voyelles dans différentes positions de la **phrase 3** dans l'expérience I (cf. chapitre IV [1.1](#))

**Modèle** : Durée ~ DAF + Voyelle + Position + DAF:Voyelle + DAF:Position + Voyelle:Position + DAF:Voyelle:Position + (1 | loc.:DAF) + (1 | loc.:Voyelle) + (1 | loc.:Position)

**Référence** : la durée moyenne des voyelles dans répétitions de phrases en Position 1 dans les essais sans DAF

	Estimée	Écart-type	DL	t-valeur	p-valeur
<b>Intercept</b>	0,112	0,006	85,540	19,392	<0,001
Position3	-0,012	0,004	21,650	-2,717	0,013
V2	-0,010	0,012	68,360	-0,790	0,432
V7	-0,032	0,012	68,360	-2,620	0,011
V9	0,081	0,012	71,570	6,522	<0,001
V2:Position3	0,017	0,009	3295,000	1,989	0,047
V9:Position2	-0,035	0,008	3295,000	-4,266	<0,001
V9:Position3	-0,068	0,009	3295,000	-7,683	<0,001
DAF60	0,019	0,005	35,050	3,689	0,001
DAF90	0,030	0,005	38,270	5,647	<0,001
DAF120	0,035	0,006	42,140	6,328	<0,001
DAF150	0,030	0,006	43,410	5,259	<0,001
DAF180	0,043	0,005	36,600	8,058	<0,001
DAF60:Position2	0,009	0,004	3319,000	2,108	0,035
DAF60:Position3	0,009	0,004	3322,000	2,219	0,027
DAF150:Position3	0,012	0,005	3247,000	2,661	0,008
DAF90:V2	0,018	0,009	3295,000	2,069	0,039
DAF180:V2	0,015	0,009	3294,000	1,754	0,079
DAF180:V2:Position3	-0,051	0,015	3295,000	-3,377	0,001
DAF90:V3	-0,023	0,009	3295,000	-2,614	0,009
DAF150:V3	-0,021	0,009	3295,000	-2,244	0,025
DAF180:V4	-0,024	0,009	3294,000	-2,764	0,006
DAF60:V5	0,036	0,008	3294,000	4,384	<0,001
DAF90:V5	0,061	0,009	3295,000	7,030	<0,001
DAF120:V5	0,062	0,009	3295,000	6,650	<0,001
DAF150:V5	0,066	0,009	3295,000	6,923	<0,001
DAF180:V5	0,078	0,009	3294,000	9,182	<0,001
DAF90:V6	-0,019	0,009	3295,000	-2,218	0,027
DAF120:V6	-0,019	0,009	3295,000	-1,997	0,046
DAF150:V6	-0,020	0,009	3295,000	-2,124	0,034
DAF180:V6	-0,023	0,009	3294,000	-2,681	0,007
DAF150:V8	0,007	0,009	3295,000	0,753	0,452
DAF150:V8:Position 2	-0,026	0,013	3293,000	-2,052	0,040
DAF150:V8:Position3	-0,029	0,013	3295,000	-2,277	0,023

	Estimée	Écart-type	DL	t-valeur	p-valeur
DAF60:V9	0,015	0,009	3295,000	1,743	0,081
DAF60:V9:Position3	0,025	0,012	3294,000	2,049	<b>0,041</b>
DAF90:V9	0,016	0,009	3296,000	1,819	0,069
DAF90:V9:Position2	0,038	0,013	3295,000	2,998	<b>0,003</b>
DAF120:V9	0,021	0,010	3296,000	2,150	<b>0,032</b>
DAF120:V9:Position2	0,042	0,013	3295,000	3,357	<b>0,001</b>
DAF120:V9:Position3	0,025	0,013	3296,000	1,961	<b>0,050</b>
DAF150:V9	<0,001	0,010	3296,000	-0,001	0,999
DAF150:V9:Position2	0,069	0,013	3294,000	5,271	<b>&lt;0,001</b>
DAF150:V9:Position3	0,049	0,013	3295,000	3,709	<b>&lt;0,001</b>
DAF180:V9	0,027	0,009	3294,000	3,026	<b>0,002</b>

Tableau 3 Résumé des résultats du modèle linéaire mixte portant sur l'effet du DAF sur la durée des voyelles dans différentes positions pour le groupe des locutrices dans la « condition uniquement DAF) dans l'expérience II (cf. chapitre IV 1.2)

*Modèle* : Durée ~ DAF + Voyelle + Phrase + DAF:Voyelle + DAF:Phrase + Voyelle:Phrase + DAF:Voyelle:Phrase + (I | loc.:DAF) + (I | loc.:Voyelle) + (I | loc.:Phrase)

*Référence* : la durée moyenne des voyelles des trois types de phrases dans les essais sans DAF

	Estimée	Écart-type	DL	t-valeur	p-valeur
<b>Intercept</b>	0,114	0,005	67,170	23,689	<b>&lt; 0,001</b>
CV	0,011	0,002	21,720	6,049	<b>&lt;0,001</b>
CCV(C)	-0,005	0,002	21,720	-2,656	<b>0,015</b>
V3	-0,030	0,008	39,940	-3,780	<b>0,001</b>
V5	0,056	0,008	39,940	7,123	<b>&lt; 0,001</b>
V2:CV	0,018	0,002	12900,000	11,661	<b>&lt; 0,001</b>
V3:CV	-0,013	0,002	12900,000	-8,355	<b>&lt; 0,001</b>
V4:CV	0,006	0,002	12900,000	4,023	<b>&lt;0,001</b>
V2:CCV(C)	0,003	0,002	12900,000	2,117	<b>0,034</b>
V3:CCV(C)	-0,004	0,002	12900,000	-2,891	<b>0,004</b>
V4:CCV(C)	-0,011	0,002	12900,000	-7,067	<b>&lt;0,001</b>
DAFf60	0,022	0,004	16,150	5,941	<b>&lt;0,001</b>
DAFf120	0,037	0,004	16,190	10,198	<b>&lt;0,001</b>
DAFf60:V2	0,004	0,002	12900,000	2,339	<b>0,019</b>
DAFf120:V2	0,010	0,002	12900,000	6,251	<b>&lt;0,001</b>
DAFf120:V3	-0,005	0,002	12900,000	-2,982	<b>0,003</b>
DAFf60:V4	-0,008	0,002	12900,000	-5,133	<b>&lt; 0,001</b>
DAFf120:V4	-0,008	0,002	12900,000	-5,292	<b>&lt; 0,001</b>
DAFf60:V5	0,015	0,002	12900,000	9,889	<b>&lt; 0,001</b>
DAFf120:V5	0,023	0,002	12900,000	14,623	<b>&lt; 0,001</b>
DAFf60:V2:CV	0,007	0,002	12900,000	3,182	<b>0,001</b>
DAFf120:V2:CV	0,014	0,002	12900,000	6,320	<b>&lt;0,001</b>
DAFf120:V4:CV	-0,005	0,002	12900,000	-2,435	<b>0,015</b>
DAFf60:V5:CV	-0,005	0,002	12900,000	-2,535	<b>0,011</b>
DAFf120:V4:CCV(C)	-0,005	0,002	12900,000	-2,227	<b>0,026</b>
DAFf60:V5:CCV(C)	0,004	0,002	12900,000	2,087	<b>0,037</b>

Tableau 4 Résumé des résultats du modèle linéaire mixte portant sur l'effet du DAF sur le ratio entre la durée totale des voyelles accentuables et non accentuables (*ratioDur*) dans l'expérience I (cf. chapitre IV 2.1)

**Modèle** :  $\text{ratioDur} \sim \text{DAF} + \text{Phrase} + \text{Position} + \text{DAF}:\text{Phrase} + \text{DAF}:\text{Position} + \text{Phrase}:\text{Position} + \text{DAF}:\text{Phrase}:\text{Position} + (1 | \text{loc.}:\text{DAF}) + (1 | \text{loc.}:\text{Position}) + (1 | \text{loc.}:\text{Phrase})$

**Référence** : PLV de la phrase 1 en position 1 dans les essais sans DAF

	Estimée	Écart-type	DL	t-valeur	p-valeur
<b>Intercept</b>	2,264	0,105	11,181	21,575	<0,001
Phrase3	-1,139	0,145	10,164	-7,865	<0,001
Position3	-0,106	0,045	15,588	-2,378	<b>0,031</b>
DAF60-DAF0	0,218	0,079	302,853	2,755	<b>0,006</b>
DAF90-DAF60	0,154	0,077	270,745	2,015	<b>0,045</b>
DAF120-DAF90	-0,032	0,071	215,041	-0,446	0,656
DAF150-DAF120	0,083	0,071	211,797	1,181	0,239
DAF180-DAF150	-0,012	0,075	242,787	-0,158	0,875
DAF180-DAF150:Position2	0,248	0,104	828,047	2,393	<b>0,017</b>
DAF180-DAF150:Phrase3:Position2	-0,503	0,156	828,687	-3,230	<b>0,001</b>

Tableau 5 Résumé des résultats des deux modèles linéaires mixtes portant sur l'effet du DAF sur le ratio entre la durée de la voyelle accentuée et celle de la voyelle non-accentuée dans chaque SA (V2/V1 et V5/V4) pour les locutrices dans la « condition uniquement DAF » dans l'expérience II, respectivement pour chaque catégorie d'analyse (cf. chapitre IV 2.2)

**Modèle** :  $\text{Ratio} \sim \text{DAF} + \text{Ratio} + \text{Phrase} + \text{DAF}:\text{Ratio} + \text{DAF}:\text{Phrase} + \text{Ratio}:\text{Phrase} + (1 | \text{loc.}:\text{DAF}) + (1 | \text{loc.}:\text{Ratio}) + (1 | \text{loc.}:\text{Phrase})$

**Référence** : le ratio V2/V1 moyen des trois types de phrases dans les essais sans DAF

Catégorie d'analyse		Estimée	Écart-type	DL	t-valeur	p-valeur
1 (phrases non finales)	<b>Intercept</b>	1,069	0,156	12,190	6,835	< 0,001
	V5/V4	1,102	0,170	4,355	6,498	<b>0,002</b>
	V5/V4:CV	-0,343	0,030	1738,000	-11,278	< 0,001
	V5/V4:CCV(C)	0,493	0,030	1738,000	16,249	< 0,001
	DAFf60	0,057	0,038	30,070	1,513	0,141
	DAFf120	0,218	0,038	30,480	5,787	< 0,001
	DAFf120:CCV(C)	-0,084	0,038	1746,000	-2,226	<b>0,026</b>
	DAFf60:V5/V4	0,088	0,052	1738,000	1,683	0,093
	DAFf120:V5/V4	-0,141	0,053	1738,000	-2,683	<b>0,007</b>
2 (phrases finales)	<b>Intercept</b>	1,178	0,097	16,350	12,129	< 0,001
	CV	0,412	0,068	18,670	6,073	< 0,001
	CCV(C)	-0,105	0,067	18,190	-1,556	0,137
	V5/V4	0,034	0,123	11,090	0,280	0,785
	V5/V4:CV	-0,499	0,029	1118,000	-17,272	< 0,001
	V5/V4:CCV(C)	0,242	0,029	1118,000	8,425	< 0,001
	DAFf60	0,071	0,035	36,160	2,005	0,053
	DAFf120	0,220	0,037	37,900	5,972	< 0,001
	DAFf60:CV	-0,073	0,035	1135,000	-2,056	<b>0,040</b>
DAFf120:V5/V4	-0,119	0,051	1120,000	-2,315	<b>0,021</b>	

Tableau 6 Résumé des résultats du modèle linéaire mixte portant sur l'effet du DAF sur la PLV dans l'expérience I (cf. chapitre V section 1)

**Modèle** :  $PLV \sim DAF + Phrase + Position + DAF:Position + Phrase:Position + (1 | loc.:DAF) + (1 | loc.:Position) + (1 | loc.:Phrase)$

**Référence** : PLV moyenne des répétitions de la phrase 1 produites en position 1 dans les essais sans DAF

	Estimée	Écart-type	DL	t-valeur	p-valeur
<b>Intercept</b>	0,498	0,023	107,651	21,928	< <b>0,001</b>
Phrase3	-0,063	0,016	16,005	-3,888	<b>0,001</b>
Phrase3:Position3	0,046	0,019	870,818	2,411	<b>0,016</b>
DAFf60	-0,017	0,026	116,979	-0,666	0,507
DAFf90	-0,017	0,026	119,526	-0,663	0,509
DAFf120	0,002	0,026	122,149	0,068	0,946
DAFf150	-0,004	0,026	127,746	-0,157	0,876
DAFf180	-0,010	0,026	121,789	-0,404	0,687
DAFf180:Position2	0,077	0,032	857,809	2,386	<b>0,017</b>
DAFf90:Position3	0,090	0,034	868,130	2,664	<b>0,008</b>
DAFf150:Position3	0,081	0,034	863,876	2,379	<b>0,018</b>
DAFf180:Position3	0,080	0,034	859,502	2,360	<b>0,019</b>

Tableau 7 Résumé des résultats du modèle linéaire mixte portant sur l'effet du DAF sur la PLV d'une phrase entière pour les locutrices dans la « condition uniquement DAF ») dans l'expérience II (cf. chapitre V 2.1)

**Modèle** :  $PLV \sim DAF + Phrase + Position + DAF:Phrase + DAF:Position + Phrase:Position + DAF:Phrase:Position + (1 | loc.:DAF) + (1 | loc.:Position) + (1 | loc.:Phrase)$

**Référence** : PLV moyenne des trois types de phrases en position 1 dans les essais sans DAF

	Estimée	Écart-type	DL	t-valeur	p-valeur
<b>Intercept</b>	0,489	0,009	78,390	53,733	< <b>0,001</b>
CCV(C)	-0,023	0,011	110,600	-2,081	<b>0,040</b>
Position2	0,023	0,011	54,010	2,169	<b>0,035</b>
Position3	0,027	0,011	54,330	2,550	<b>0,014</b>
CV:Position3	0,028	0,012	2623,000	2,340	<b>0,019</b>
DAFf60	0,028	0,009	112,400	3,161	<b>0,002</b>
DAFf120	0,029	0,009	113,600	3,214	<b>0,002</b>
DAF60:Position2	-0,031	0,012	2599,000	-2,588	<b>0,010</b>
DAF60:Position3	-0,020	0,012	2604,000	-1,630	0,103
DAF120:Position2	-0,017	0,012	2602,000	-1,406	0,160
DAF120:Position3	-0,013	0,012	2604,000	-1,095	0,274
DAF120:CCV(C):Position3	-0,034	0,017	2634,000	-2,013	<b>0,044</b>

Tableau 8 Résumé des résultats des deux modèles linéaires mixtes portant sur l'effet du DAF sur la PLV de chaque SA pour les locutrices dans la « condition uniquement DAF » dans l'expérience II, respectivement pour chaque catégorie d'analyse (cf. chapitre V 2.2)

Modèle :  $PLV \sim DAF + Phrase + Syntagme + DAF:Phrase + DAF:Syntagme + Phrase:Syntagme + DAF:Phrase:Syntagme + (I | loc.:DAF) + (I | loc.:Phrase) + (I | loc.:Syntagme)$

Référence : PLV moyenne du SA initial des trois types de phrases dans les essais sans DAF

Catégorie d'analyse		Estimée	Écart-type	DL	t-valeur	p-valeur
1 (phrases non finales)	Intercept	0,700	0,032	11,020	21,795	< 0,001
	CCV(C)	-0,071	0,017	203,200	-4,067	< 0,001
	SAfinal	0,005	0,041	6,783	0,124	0,905
	CCV(C):SAfinal	0,045	0,024	1759,000	1,866	0,062
	DAFf60	0,036	0,025	10,390	1,444	0,178
	DAFf120	0,040	0,025	10,370	1,597	0,140
	DAFf60:CCV(C)	0,048	0,024	1767,000	2,035	0,042
	DAFf120:CCV(C)	0,070	0,024	1767,000	2,909	0,004
	DAFf60:SAfinal	0,054	0,024	1759,000	2,260	0,024
	DAFf120:SAfinal	0,056	0,024	1759,000	2,352	0,019
DAFf120:CCV(C):SAfinal	-0,082	0,034	1760,000	-2,392	0,017	
2 (phrases finales)	Intercept	0,722	0,027	22,230	26,490	< 0,001
	SAfinal	-0,043	0,036	14,950	-1,180	0,256
	CV: SAfinal	-0,100	0,032	1208,000	-3,126	0,002
	CCV(C): SAfinal	0,065	0,031	1210,000	2,131	0,033
	DAFf60	0,018	0,026	21,930	0,693	0,495
	DAFf120	0,052	0,026	21,940	2,015	0,056
	DAFf60:SAfinal	0,026	0,031	1206,000	0,837	0,403
	DAFf120:SAfinal	0,011	0,031	1206,000	0,346	0,729
	DAFf120:CV:SAfinal	0,121	0,045	1209,000	2,705	0,007

Tableau 9 Récapitulatif des 20 modèles linéaires portant sur l'effet décalFO, conduits individuellement pour chaque locutrice (ex. Loc01 = locutrice 01) ayant subi le décalFO dans l'expérience II (cf. chapitre VI section 1)

GROUPE1 (décalage positif)		GROUPE2 (décalage négatif)	
Locutrice	Modèle	Locutrice	Modèle
Loc01	f0 ~ décalage <sup>16</sup> + phrase + accent <sup>17</sup> + 1	Loc11	f0 ~ décalage + accent + 1
Loc02	f0 ~ décalage + phrase + accent + phrase:accent + 1	Loc12	f0 ~ décalage + accent + 1
Loc03	f0 ~ décalage + accent + décalage:accent + 1	Loc13	f0 ~ décalage + phrase + accent + phrase:accent + 1
Loc04	f0 ~ décalage + accent + 1	Loc14	f0 ~ décalage + phrase + accent + 1
Loc05	f0 ~ décalage + accent + 1	Loc15	f0 ~ décalage + phrase + accent + décalage:accent + phrase:accent + 1

<sup>16</sup> La variable « décalage » fait référence au niveau du décalage de la f0 (décalage0 = les essais de base ; décalage 1 = les essais avec un demi-ton du décalage ; décalage 2 = les essais avec deux demi-tons du décalage).

<sup>17</sup> La variable « accent » fait référence au type de voyelle, englobant les voyelles non accentuables et les voyelles accentuables.

GROUPE1 (décalage positif)		GROUPE2 (décalage négatif)	
Locutrice	Modèle	Locutrice	Modèle
Loc06	f0 ~ décalage + accent + décalage:accent + 1	Loc16	f0 ~ décalage + 1
Loc07	f0 ~ décalage + phrase + accent + décalage:sentID + 1	Loc17	f0 ~ décalage + phrase + accent + 1
Loc08	f0 ~ décalage + phrase + accent + phrase:accent + 1	Loc18	f0 ~ décalage + phrase + 1
Loc09	f0 ~ décalage + accent + 1	Loc19	f0 ~ décalage + accent + décalage:accent + 1
Loc10	f0 ~ décalage + 1	Loc20	f0 ~ décalage + phrase + accent + 1

Tableau 10 Résumé des résultats des 20 modèles linéaires portant sur l'effet du décalF0, conduits individuellement pour chaque locutrice ayant subi le décalF0 dans l'expérience II (cf. chapitre VI section 1)

Groupe	Locutrice		Intercept	Décalage1	Décalage2	Décalage1:Accent	Décalage2:Accent
1 (Groupe positif)	01	Estimée	237,798	3,652	3,847	/	/
		p.valeur	<0,001	<b>0,017</b>	<b>0,015</b>	/	/
	02	Estimée	231,283	5,821	5,867	/	/
		p.valeur	<0,001	0,096	0,093	/	/
	03	Estimée	220,945	-3,638	-7,424	/	/
		p.valeur	<0,001	0,126	<b>0,002</b>	/	/
	04	Estimée	214,40	25,220	24,83	/	/
		p.valeur	<0,001	<b>&lt;0,001</b>	<b>&lt;0,001</b>	/	/
	05	Estimée	233,012	3,197	8,339	/	/
		p.valeur	<0,001	0,423	<b>0,037</b>	/	/
06	Estimée	165,48437	11,630	16,070	13,639	13,976	
	p.valeur	<0,001	<b>0,013</b>	<b>&lt;0,001</b>	<b>0,029</b>	<b>0,025</b>	
07	Estimée	234,702	16,796	16,538	/	/	
	p.valeur	<0,001	<b>&lt;0,001</b>	<b>&lt;0,001</b>	/	/	
08	Estimée	220,845	0,105	2,883	/	/	
	p.valeur	<0,001	0,980	0,489	/	/	
09	Estimée	207,866	-7,722	-8,779	/	/	
	p.valeur	<0,001	<b>0,004</b>	<b>0,001</b>	/	/	
10	Estimée	188,492	9,201	17,378	/	/	
	p.valeur	<0,001	<b>0,003</b>	<b>&lt;0,001</b>	/	/	
2 (Groupe négatif)	11	Estimée	178,450	9,466	12,000	/	/
		p.valeur	<0,001	<b>0,002</b>	<b>&lt;0,001</b>	/	/
	12	Estimée	232,966	-9,820	-5,804	/	/
		p.valeur	<0,001	<b>0,011</b>	0,132	/	/
	13	Estimée	237,650	-18,250	-13,040	/	/
		p.valeur	<0,001	<b>&lt;0,001</b>	<b>&lt;0,001</b>	/	/
	14	Estimée	190,626	-1,177	-4,037	/	/
		p.valeur	<0,001	0,698	0,183	/	/
	15	Estimée	235,045	-10,695	-8,286	/	/
		p.valeur	<0,001	<b>0,004</b>	<b>0,025</b>	/	/
	16	Estimée	166,537	-0,191	-3,046	/	/
		p.valeur	<0,001	0,926	0,139	/	/

Groupe	Locutrice	Intercept	Décalage1	Décalage2	Décalage1:Accent	Décalage2:Accent
17	Estimée	212,429	-0,622	-8,327	/	/
	p.valeur	<0,001	0,865	<b>0,023</b>	/	/
18	Estimée	192,296	-12,915	-7,409	/	/
	p.valeur	<0,001	<b>&lt;0,001</b>	<b>0,006</b>	/	/
19	Estimée	208,053	-9,336	-5,844	-11,192	-12,093
	p.valeur	<0,001	<b>0,019</b>	0,141	<b>0,036</b>	<b>0,024</b>
20	Estimée	190,249	3,414	3,166	/	/
	p.valeur	<0,001	0,184	0,633	/	/

Tableau 11 Résumé des résultats des 12 modèles linéaires mixtes comparant l'effet du DAF sur la durée des voyelles dans chaque SA de chaque type de phrase entre les différentes conditions du décalFO pour les locutrices ayant suivi le décalage dans l'expérience II (cf. chapitre VI 2.1.1)

**Modèle** : Durée ~ décalage + DAF + Voyelle + décalage:DAF + décalage:Voyelle + DAF :Voyelle + décalage:DAF :Voyelle + (I | loc.:décalage) + (I | locc.:DAF) + (I | loc.:Voyelle)

**Référence** : la durée de la voyelle non accentuée (V1 dans SA initial et V4 dans SA final) dans chaque SA dans les essais sans décalFO et DAF

Groupe	Phrase-SA	Estimée	Écart-type	DL	t-valeur	p-valeur
CV-initial (1406 obs.)	<b>Intercept</b>	0,123	0,012	17,270	10,378	<b>&lt;0,001</b>
	DAF60	0,008	0,005	65,090	1,474	0,145
	DAF120	0,013	0,005	66,210	2,480	<b>0,016</b>
	DAF60:V2	0,013	0,006	1346,000	2,230	<b>0,026</b>
	DAF120:V2	0,029	0,006	1346,000	4,910	<b>&lt;0,001</b>
	Décal.1:DAF60	0,008	0,006	1346,000	1,313	0,189
	Décal.2:DAF60	0,005	0,006	1346,000	0,917	0,359
	Décal.1:DAF120	0,009	0,006	1347,000	1,527	0,127
	Décal.2:DAF120	0,003	0,006	1347,000	0,495	0,621
	Décal.1:DAF60:V2	-0,001	0,008	1346,000	-0,139	0,889
	Décal.2:DAF60:V2	0,001	0,008	1346,000	0,152	0,880
	Décal.1:DAF120:V2	-0,003	0,008	1346,000	-0,404	0,686
	Décal.2:DAF120:V2	0,004	0,008	1346,000	0,503	0,615
Groupe positif CVC-initial (1422 obs.)	<b>Intercept</b>	0,113	0,006	23,070	19,388	<b>&lt;0,001</b>
	V2	-0,031	0,008	17,320	-4,091	<b>0,001</b>
	DAF60	0,017	0,004	85,740	4,363	<b>&lt;0,001</b>
	DAF120	0,026	0,004	86,510	6,675	<b>&lt;0,001</b>
	DAF60:V2	<0,001	0,005	1362,000	-0,005	0,996
	DAF120:V2	-0,003	0,005	1362,000	-0,755	0,451
	Décal.1:DAF60	0,003	0,005	1362,000	0,572	0,567
	Décal.2:DAF60	-0,004	0,005	1362,000	-0,791	0,429
	Décal.1:DAF120	0,001	0,005	1362,000	0,172	0,863
	Décal.2:DAF120	-0,007	0,005	1362,000	-1,425	0,154
	Décal.1:DAF60:V2	0,001	0,007	1362,000	0,228	0,820
	Décal.2:DAF60:V2	-0,001	0,007	1362,000	-0,085	0,932
	Décal.1:DAF120:V2	0,004	0,007	1362,000	0,649	0,516
<b>Décal.2:DAF120:V2</b>	<b>0,013</b>	<b>0,007</b>	<b>1362,000</b>	<b>1,985</b>	<b>0,047</b>	
CCV-initial (1422 obs.)	<b>Intercept</b>	0,100	0,014	16,080	7,190	<b>&lt;0,001</b>
	DAF60	0,015	0,005	84,680	3,010	<b>0,003</b>
	DAF120	0,017	0,005	85,400	3,326	<b>0,001</b>

Groupe	Phrase-SA	Estimée	Écart-type	DL	t-valeur	p-valeur
	DAF60:V2	-0,003	0,006	1362,000	-0,514	0,607
	DAF120:V2	0,017	0,006	1362,000	2,893	<b>0,004</b>
	Décal.1:DAF60	-0,003	0,006	1362,000	-0,531	0,595
	Décal.2:DAF60	-0,003	0,006	1362,000	-0,561	0,575
	Décal.1:DAF120	-0,005	0,006	1362,000	-0,812	0,417
	Décal.2:DAF120	-0,001	0,006	1362,000	-0,240	0,811
	Décal.1:DAF60:V2	0,010	0,008	1362,000	1,225	0,221
	Décal.2:DAF60:V2	0,012	0,008	1362,000	1,460	0,145
	Décal.1:DAF120:V2	0,009	0,008	1362,000	1,042	0,298
	Décal.2:DAF120:V2	0,007	0,008	1362,000	0,860	0,390
	<b>Intercept</b>	0,113	0,012	20,990	9,042	<b>&lt;0,001</b>
	V5	0,125	0,017	17,910	7,394	<b>&lt;0,001</b>
	DAF60	0,014	0,007	127,300	1,847	0,067
	DAF120	0,016	0,008	135,500	2,140	<b>0,034</b>
	DAF60:V5	0,057	0,010	894,600	5,945	<b>&lt;0,001</b>
	DAF120:V5	0,074	0,010	894,700	7,570	<b>&lt;0,001</b>
CV-final (954 obs.)	Décal.1:DAF60	0,000	0,010	897,300	-0,007	0,994
	Décal.2:DAF60	-0,001	0,010	904,000	-0,071	0,944
	Décal.1:DAF120	0,002	0,010	900,200	0,197	0,844
	Décal.2:DAF120	0,006	0,010	903,100	0,574	0,566
	<b>Décal.1:DAF60:V5</b>	<b>-0,032</b>	<b>0,014</b>	<b>894,700</b>	<b>-2,346</b>	<b>0,019</b>
	<b>Décal.2:DAF60:V5</b>	<b>-0,039</b>	<b>0,014</b>	<b>894,700</b>	<b>-2,876</b>	<b>0,004</b>
	Décal.1:DAF120:V5	-0,019	0,014	894,700	-1,389	0,165
	<b>Décal.2:DAF120 :V5</b>	<b>-0,040</b>	<b>0,014</b>	<b>894,900</b>	<b>-2,958</b>	<b>0,003</b>
	<b>Intercept</b>	0,115	0,013	23,733	8,940	<b>&lt;0,001</b>
	V5	0,075	0,017	18,791	4,397	<b>&lt;0,001</b>
	DAF60	0,009	0,010	65,233	0,929	0,356
	DAF120	0,027	0,010	63,461	2,815	<b>0,006</b>
	DAF60:V5	0,035	0,011	899,437	3,206	<b>0,001</b>
	DAF120:V5	0,044	0,011	899,320	4,064	<b>&lt;0,001</b>
CVC-final (958 obs.)	Décal.1:DAF60	0,004	0,011	909,177	0,319	0,750
	Décal.2:DAF60	-0,001	0,011	908,592	-0,128	0,898
	Décal.1:DAF120	0,005	0,011	906,423	0,466	0,641
	Décal.2:DAF120	0,003	0,011	912,178	0,272	0,785
	Décal.1:DAF60:V5	-0,029	0,016	899,564	-1,825	0,068
	Décal.2:DAF60:V5	-0,020	0,015	899,451	-1,318	0,188
	Décal.1:DAF120:V5	-0,010	0,015	899,368	-0,680	0,497
	<b>Décal.2:DAF120:V5</b>	<b>-0,033</b>	<b>0,016</b>	<b>899,494</b>	<b>-2,084</b>	<b>0,037</b>
	<b>Intercept</b>	0,084	0,011	27,080	7,810	<b>&lt;0,001</b>
	V5	0,135	0,013	17,130	10,004	<b>&lt;0,001</b>
	V5:Décal.2	0,019	0,009	894,700	2,095	<b>0,036</b>
	DAF60	0,008	0,009	40,460	0,894	0,377
	DAF120	0,021	0,009	39,050	2,308	<b>0,026</b>
	DAF60:V5	0,033	0,009	895,300	3,636	<b>&lt;0,001</b>
	DAF120:V5	0,063	0,009	894,700	7,119	<b>&lt;0,001</b>
CCV-final (954 obs.)	Décal.1:DAF60	0,003	0,009	903,700	0,298	0,765
	Décal.2:DAF60	0,007	0,009	903,300	0,809	0,419
	Décal.1:DAF120	-0,004	0,009	899,500	-0,391	0,696
	Décal.2:DAF120	-0,002	0,009	899,400	-0,256	0,798
	Décal.1:DAF60:V5	-0,019	0,013	895,400	-1,526	0,127
	<b>Décal.2:DAF60:V5</b>	<b>-0,038</b>	<b>0,013</b>	<b>895,000</b>	<b>-2,940</b>	<b>0,003</b>
	<b>Décal.1:DAF120:V5</b>	<b>-0,034</b>	<b>0,013</b>	<b>894,900</b>	<b>-2,683</b>	<b>0,007</b>

Groupe	Phrase-SA	Estimée	Écart-type	DL	t-valeur	p-valeur	
	<b>Décal.2:DAF120:V5</b>	<b>-0,042</b>	<b>0,013</b>	<b>894,800</b>	<b>-3,256</b>	<b>0,001</b>	
CV-initial (1198 obs.)	<b>Intercept</b>	0,105	0,009	17,050	11,850	<b>&lt;0,001</b>	
	DAF60	0,014	0,005	70,320	2,706	<b>0,009</b>	
	DAF120	0,020	0,005	77,570	3,634	<b>0,001</b>	
	DAF60:V2	0,011	0,006	1144,000	1,696	<b>0,090</b>	
	DAF120:V2	0,028	0,007	1145,000	4,328	<b>&lt;0,001</b>	
	Décal.1:DAF60	-0,001	0,006	1145,000	-0,218	0,828	
	Décal.2:DAF60	0,003	0,006	1145,000	0,495	0,621	
	Décal.1:DAF120	-0,004	0,006	1149,000	-0,558	0,577	
	Décal.2:DAF120	<0,001	0,006	1148,000	-0,075	0,940	
	Décal.1:DAF60:V2	<0,001	0,009	1144,000	0,037	0,970	
	Décal.2:DAF60:V2	-0,004	0,009	1144,000	-0,468	0,640	
	Décal.1:DAF120:V2	0,007	0,009	1145,000	0,715	0,474	
	Décal.2:DAF120:V2	-0,004	0,009	1145,000	-0,421	0,674	
	CVC-initial (1226 obs.)	<b>Intercept</b>	0,118	0,008	23,160	14,824	<b>&lt;0,001</b>
		V2	-0,050	0,010	13,480	-5,226	<b>&lt;0,001</b>
DAF60		0,018	0,007	23,370	2,598	<b>0,016</b>	
DAF120		0,024	0,007	23,730	3,481	<b>0,002</b>	
DAF60:V2		0,004	0,005	1172,000	0,773	0,440	
DAF120:V2		0,016	0,006	1172,000	2,811	<b>0,005</b>	
Décal.1:DAF60		-0,005	0,005	1172,000	-0,824	0,410	
Décal.2:DAF60		-0,004	0,006	1172,000	-0,804	0,422	
Décal.1:DAF120		-0,007	0,006	1172,000	-1,285	0,199	
Décal.2:DAF120		-0,009	0,006	1172,000	-1,567	0,117	
Décal.1:DAF60:V2		<0,001	0,008	1172,000	-0,019	0,985	
Décal.2:DAF60:V2		0,006	0,008	1172,000	0,799	0,425	
Décal.1:DAF120:V2		0,004	0,008	1172,000	0,545	0,586	
Décal.2:DAF120:V2	0,009	0,008	1172,000	1,176	0,240		
CCV-initial (1212 obs.)	<b>Intercept</b>	0,097	0,014	14,340	6,906	<b>&lt;0,001</b>	
	DAF60	0,012	0,007	28,770	1,879	<b>0,070</b>	
	DAF120	0,016	0,007	28,770	2,475	<b>0,020</b>	
	DAF60:V2	0,009	0,006	1158,000	1,501	0,134	
	DAF120:V2	0,030	0,006	1158,000	4,947	<b>&lt;0,001</b>	
	Décal.1:DAF60	-0,009	0,006	1158,000	-1,402	0,161	
	Décal.2:DAF60	-0,001	0,006	1158,000	-0,134	0,893	
	<b>Décal.1:DAF120</b>	<b>-0,011</b>	<b>0,006</b>	<b>1159,000</b>	<b>-1,745</b>	<b>0,081</b>	
	Décal.2:DAF120	-0,005	0,006	1158,000	-0,743	0,457	
	Décal.1:DAF60:V2	0,011	0,009	1158,000	1,310	0,190	
	Décal.2:DAF60:V2	0,007	0,009	1158,000	0,861	0,390	
	Décal.1:DAF120:V2	0,008	0,009	1158,000	0,896	0,371	
	Décal.2:DAF120:V2	0,003	0,009	1158,000	0,298	0,766	
CV-final (804 obs.)	<b>Intercept</b>	0,117	0,009	30,814	12,640	<b>&lt;0,001</b>	
	V5	0,089	0,013	24,075	7,140	<b>&lt;0,001</b>	
	DAF60	0,025	0,009	93,520	2,871	<b>0,005</b>	
	DAF120	0,026	0,009	103,905	2,873	<b>0,005</b>	
	DAF60:V5	0,024	0,011	750,106	2,123	<b>0,034</b>	
	DAF120:V5	0,049	0,012	749,778	4,219	<b>&lt;0,001</b>	
	Décal.1:DAF60	-0,015	0,011	758,296	-1,324	0,186	
	Décal.2:DAF60	-0,011	0,011	762,158	-0,999	0,318	
	Décal.1:DAF120	-0,003	0,012	760,275	-0,238	0,812	
	Décal.2:DAF120	-0,008	0,011	762,933	-0,684	0,494	
	Décal.1:DAF60:V5	0,002	0,016	750,204	0,101	0,919	

Groupe	Phrase-SA	Estimée	Écart-type	DL	t-valeur	p-valeur
	Décal.2:DAF60:V5	0,002	0,016	750,693	0,155	0,877
	Décal.1:DAF120:V5	-0,005	0,016	749,962	-0,320	0,749
	Décal.2:DAF120:V5	0,007	0,016	749,776	0,407	0,684
	<b>Intercept</b>	0,099	0,015	15,994	6,585	<b>&lt;0,001</b>
	V5	0,055	0,020	13,306	2,724	<b>0,017</b>
	DAF60	0,021	0,009	41,027	2,312	<b>0,026</b>
	DAF120	0,039	0,009	42,158	4,348	<b>&lt;0,001</b>
	DAF60:V5	0,016	0,009	774,623	1,683	0,093
	DAF120:V5	0,029	0,010	774,637	3,011	<b>0,003</b>
CVC-final (828 obs.)	Décal.1:DAF60	0,002	0,009	778,398	0,258	0,796
	Décal.2:DAF60	0,009	0,010	778,939	0,958	0,338
	Décal.1:DAF120	-0,004	0,010	777,478	-0,382	0,703
	Décal.2:DAF120	0,004	0,010	776,994	0,418	0,676
	Décal.1:DAF60:V5	-0,013	0,013	774,684	-0,985	0,325
	Décal.2:DAF60:V5	-0,015	0,013	774,723	-1,089	0,277
	Décal.1:DAF120:V5	-0,008	0,014	774,703	-0,588	0,557
	Décal.2:DAF120:V5	-0,014	0,014	774,621	-1,004	0,315
	<b>Intercept</b>	0,084	0,010	23,120	8,295	<b>&lt;0,001</b>
	V5	0,100	0,013	16,420	7,709	<b>&lt;0,001</b>
	DAF60	0,013	0,008	45,380	1,563	0,125
	DAF120	0,019	0,008	40,810	2,324	<b>0,025</b>
	DAF60:V5	0,033	0,009	758,400	3,674	<b>&lt;0,001</b>
	DAF120:V5	0,040	0,009	758,200	4,726	<b>&lt;0,001</b>
CCV-final (812 obs.)	Décal.1:DAF60	-0,001	0,009	763,900	-0,067	0,946
	Décal.2:DAF60	<0,001	0,009	762,000	-0,052	0,958
	Décal.1:DAF120	0,004	0,009	761,300	0,475	0,635
	Décal.2:DAF120	-0,006	0,009	759,700	-0,644	0,520
	Décal.1:DAF60:V5	-0,009	0,013	758,700	-0,734	0,463
	Décal.2:DAF60:V5	-0,016	0,013	758,300	-1,279	0,201
	Décal.1:DAF120:V5	0,004	0,012	758,300	0,295	0,768
	Décal.2:DAF120:V5	-0,012	0,012	758,300	-0,972	0,332

Tableau 12 Résumé des résultats des six modèles linéaires mixtes comparant l'effet du DAF sur la durée des voyelles dans chaque SA de chaque type de phrase entre les locutrices dans la « condition uniquement DAF » et les locutrices dans la « condition perturbations multiples » dans l'expérience II (cf. chapitre VI 2.1.2)

**Modèle :**  $Durée \sim Groupe + DAF + Voyelle + Groupe:DAF + Groupe:Voyelle + DAF:Voyelle + Groupe:DAF:Voyelle + (1 | loc.:DAF) + (1 | loc.:Voyelle)$

**Référence :** la durée de la voyelle non accentuée (V1 dans SA initial et V4 dans SA final) pour le groupe de locutrices dans la « condition uniquement DAF » dans les essai sans DAF

Groupe1 = locutrices ayant subi DAF et décalF0 positif ; Groupe2 = locutrices ayant subi DAF et décalF0 négatif

Phrase-SA	Estimée	Écart-type	DL	t-valeur	p-valeur	
	<b>Intercept</b>	0,099	0,009	47,670	10,817	<b>&lt;0,001</b>
	V2	0,040	0,013	44,100	3,170	<b>0,003</b>
CV-initial (3472 obs.)	DAF60	0,014	0,003	59,820	4,215	<b>&lt;0,001</b>
	DAF120	0,016	0,003	60,660	4,863	<b>&lt;0,001</b>
	DAF60:V2	0,018	0,003	3366,000	6,077	<b>&lt;0,001</b>

Phrase-SA		Estimée	Écart-type	DL	t-valeur	p-valeur
	DAF120:V2	0,044	0,003	3366,000	14,661	<0,001
	Groupe1:DAF60:V2	-0,005	0,005	3366,000	-0,997	0,319
	Groupe2:DAF60:V2	-0,009	0,005	3366,000	-1,771	0,077
	<b>Groupe1:DAF120:V2</b>	<b>-0,015</b>	<b>0,005</b>	<b>3366,000</b>	<b>-2,847</b>	<b>0,004</b>
	<b>Groupe2:DAF120:V2</b>	<b>-0,015</b>	<b>0,005</b>	<b>3366,000</b>	<b>-2,755</b>	<b>0,006</b>
	<b>Intercept</b>	0,101	0,006	52,060	16,049	<0,001
	V2	-0,017	0,008	41,500	-2,044	0,047
	DAF60	0,013	0,003	50,080	3,867	<0,001
	DAF120	0,017	0,003	51,600	4,825	<0,001
CVC-initial (3486 obs.)	DAF60:V2	0,005	0,002	3379,000	2,129	0,033
	DAF120:V2	0,017	0,003	3380,000	6,806	<0,001
	Groupe1:DAF60:V2	-0,005	0,004	3379,000	-1,146	0,252
	Groupe2:DAF60:V2	0,002	0,004	3379,000	0,434	0,664
	<b>Groupe1:DAF120:V2</b>	<b>-0,012</b>	<b>0,004</b>	<b>3379,000</b>	<b>-2,847</b>	<b>0,004</b>
	Groupe2:DAF120:V2	0,005	0,004	3379,000	1,111	0,267
	<b>Intercept</b>	0,107	0,011	47,450	9,481	<0,001
	DAF60	0,012	0,004	52,570	2,873	0,006
	DAF120	0,019	0,004	52,940	4,747	<0,001
	DAF60:V2	0,014	0,003	3402,000	4,901	<0,001
	DAF120:V2	0,026	0,003	3402,000	9,114	<0,001
CCV-initial (3508 obs.)	Groupe1:DAF60:V2	-0,006	0,005	3402,000	-1,153	0,249
	Groupe2:DAF60:V2	0,005	0,005	3402,000	0,910	0,363
	Groupe1:DAF120:V2	<0,001	0,005	3402,000	-0,085	0,932
	Groupe2:DAF120:V2	0,009	0,005	3402,000	1,840	0,066
	<b>Intercept</b>	0,121	0,013	49,560	9,163	<0,001
	V5	0,089	0,018	44,470	4,928	<0,001
	DAF60	0,010	0,006	63,390	1,780	0,080
	DAF120	0,022	0,006	63,680	3,843	<0,001
CV-final (2338 obs.)	DAF60:V5	0,024	0,005	2234,000	4,678	<0,001
	DAF120:V5	0,036	0,005	2234,000	6,972	<0,001
	Groupe1:DAF60:V5	-0,003	0,009	2235,000	-0,371	0,711
	Groupe2:DAF60:V5	0,001	0,009	2235,000	0,141	0,888
	Groupe1:DAF120:V5	0,008	0,009	2235,000	0,885	0,376
	Groupe2:DAF120:V5	0,014	0,009	2235,000	1,460	0,144
	<b>Intercept</b>	0,102	0,013	51,330	7,687	<0,001
	V5	0,071	0,018	42,660	3,972	<0,001
	DAF60	0,018	0,007	54,080	2,725	0,009
	DAF120	0,035	0,007	54,500	5,308	<0,001
CVC-final (2366 obs.)	DAF60:V5	0,022	0,005	2262,000	4,414	<0,001
	DAF120:V5	0,035	0,005	2262,000	7,120	<0,001
	Groupe1:DAF60:V5	-0,009	0,008	2262,000	-1,038	0,300
	<b>Groupe2:DAF60:V5</b>	<b>-0,020</b>	<b>0,009</b>	<b>2262,000</b>	<b>-2,295</b>	<b>0,022</b>
	Groupe1:DAF120:V5	-0,010	0,008	2262,000	-1,271	0,204
	<b>Groupe2:DAF120:V5</b>	<b>-0,017</b>	<b>0,009</b>	<b>2261,000</b>	<b>-1,980</b>	<b>0,048</b>
	<b>Intercept</b>	0,088	0,011	55,280	8,194	<0,001
	V5	0,114	0,014	43,220	8,014	<0,001
	DAF60	0,008	0,006	55,180	1,270	0,209
	DAF120	0,021	0,006	55,290	3,365	0,001
CCV-final (2354 obs.)	DAF60:V5	0,025	0,005	2249,000	5,346	<0,001
	DAF120:V5	0,026	0,005	2249,000	5,581	<0,001
	<b>Groupe1:DAF60:V5</b>	<b>-0,020</b>	<b>0,008</b>	<b>2249,000</b>	<b>-2,619</b>	<b>0,009</b>
	Groupe2:DAF60:V5	-0,005	0,008	2249,000	-0,597	0,551

Phrase-SA	Estimée	Écart-type	DL	t-valeur	p-valeur
Groupe1:DAF120:V5	<0,001	0,008	2250,000	-0,051	0,959
Groupe2:DAF120:V5	0,010	0,008	2249,000	1,240	0,215

Tableau 13 Résumé des résultats des six modèles linéaires mixtes comparant l'effet du DAF sur la PLV (d'une phrase entière ou de chaque SA) entre les différentes conditions du décalFO dans l'expérience II (cf. chapitre VI 2.2.1)

**Modèle** :  $PLV \sim \text{Décalage} + \text{DAF} + \text{Phrase} + \text{Décalage:DAF} + \text{Décalage:Phrase} + \text{DAF:Phrase} + \text{Décalage:DAF:Phrase} + (1 | \text{loc.}:\text{Décalage}) + (1 | \text{loc.}:\text{DAF}) + (1 | \text{loc.}:\text{Phrase})$

**Référence** : la PLV moyenne des trois types de phrases dans les essais sans décalFO et DAF

Décal.1 = 1 demi-ton du décalFO ; Décal.2 = 2 demi-tons du décalFO

Groupe	Intervalle		Estimée	Écart-type	DL	t-valeur	p-valeur	
Groupe positif	Phrase	<b>Intercept</b>	0,482	0,013	113,900	37,908	<b>&lt;0,001</b>	
		<b>Intercept</b>	0,648	0,032	68,570	20,101	<b>&lt;0,001</b>	
	SA initial	DAF60	0,096	0,041	55,180	2,340	<b>0,023</b>	
		DAF120	0,080	0,041	55,310	1,952	0,056	
		DAF120:CV	0,049	0,050	666,700	0,999	0,318	
		<b>Décal.1:DAF120:CV</b>	<b>-0,142</b>	<b>0,070</b>	<b>673,100</b>	<b>-2,033</b>	<b>0,042</b>	
	SA final	<b>Intercept</b>	0,759	0,023	92,879	33,114	<b>&lt;0,001</b>	
	Groupe négatif	Phrase	<b>Intercept</b>	0,497	0,018	53,164	27,700	<b>&lt;0,001</b>
			<b>Intercept</b>	0,735	0,034	59,030	21,732	<b>&lt;0,001</b>
		SA initial	Décal.1	-0,030	0,039	52,120	-0,759	0,451
Décal.2			-0,013	0,039	51,120	-0,343	0,733	
DAF60			0,044	0,039	57,090	1,153	0,254	
DAF120			0,002	0,039	58,080	0,059	0,953	
<b>Décal.1:DAF120</b>			<b>0,107</b>	<b>0,050</b>	<b>566,800</b>	<b>2,139</b>	<b>0,033</b>	
SA final		<b>Intercept</b>	0,787	0,027	56,820	28,983	<b>&lt;0,001</b>	
		Décal.1	-0,090	0,030	68,150	-2,973	<b>0,004</b>	
		Décal.2	-0,041	0,030	66,740	-1,355	0,180	
	DAF60	0,048	0,034	42,040	1,422	0,162		
	DAF120	0,035	0,034	42,700	1,046	0,301		
	Décal.1:DAF60	0,023	0,041	571,400	0,557	0,578		
	Décal.2:DAF60	0,044	0,041	569,800	1,088	0,277		
	<b>Décal.1:DAF120</b>	<b>0,109</b>	<b>0,041</b>	<b>568,400</b>	<b>2,677</b>	<b>0,008</b>		
<b>Décal.2:DAF120</b>	<b>0,093</b>	<b>0,041</b>	<b>566,300</b>	<b>2,300</b>	<b>0,022</b>			

Tableau 14 Résumé des résultats des trois modèles linéaires mixtes comparant l'effet du DAF sur la PLV (d'une phrase entière ou de chaque SA) entre les locutrices dans la « condition uniquement

DAF » et les locutrices dans la « condition perturbations multiples » dans l'expérience II (cf. chapitre VI 2.2.2)

**Modèle :**  $PLV \sim \text{Groupe} + \text{DAF} + \text{Phrase} + \text{Groupe:DAF} + \text{Groupe:Phrase} + \text{DAF:Phrase} + \text{Groupe:DAF:Phrase} + (1 | \text{loc.:Phrase}) + (1 | \text{loc.:DAF})$

**Référence :** PLV moyenne des trois types de phrases dans les essais sans DAF pour les locutrices dans la « condition uniquement DAF »

Groupe1 = Groupe Positif ; Groupe2 = Groupe Négatif

Intervalle		Estimée	Écart-type	DL	t-valeur	p-valeur
Phrase	<b>Intercept</b>	0,491	0,009	75,490	52,280	<0,001
	DAF60	0,025	0,009	29,220	2,902	0,007
	DAF120	0,026	0,009	29,610	2,952	0,006
SA initial	<b>Intercept</b>	0,669	0,023	71,190	28,540	<0,001
	Groupe2	0,043	0,038	82,820	1,146	0,255
	Groupe2:CCV	0,088	0,043	105,800	2,036	0,044
	DAF60	0,067	0,027	27,740	2,493	0,019
	DAF60:CCV	-0,006	0,025	1706,000	-0,245	0,806
	<b>Groupe2:DAF60:CCV</b>	<b>-0,098</b>	<b>0,044</b>	<b>1706,000</b>	<b>-2,222</b>	<b>0,026</b>
	DAF120	0,080	0,027	27,930	2,993	0,006
	DAF120:CV	0,033	0,025	1692,000	1,356	0,175
<b>Groupe1:DAF120:CV</b>	<b>-0,108</b>	<b>0,042</b>	<b>1716,000</b>	<b>-2,571</b>	<b>0,010</b>	
SA final	<b>Intercept</b>	0,714	0,018	65,940	38,959	<0,001
	DAF60	0,062	0,022	26,400	2,825	0,009
	DAF120	0,094	0,022	26,600	4,289	<0,001
	DAF120:CV	0,033	0,021	1704,000	1,537	0,125
	<b>Groupe1:DAF120:CV</b>	<b>-0,085</b>	<b>0,036</b>	<b>1728,000</b>	<b>-2,331</b>	<b>0,020</b>

Tableau 15 Résumé des résultats des six modèles linéaires mixtes comparant l'effet du DAF sur la f0 médiane des voyelles dans chaque SA (initial ou final) entre le groupe des locutrices dans la « condition uniquement DAF » et les locutrices dans la « condition perturbations multiples » dans l'expérience II (cf. chapitre VI 2.3.1)

**Modèle :**  $F0 \sim \text{Groupe} + \text{DAF} + \text{Voyelle} + \text{Groupe:DAF} + \text{Groupe:Voyelle} + \text{DAF:Voyelle} + \text{Groupe:DAF:Voyelle} + (1 | \text{loc.:Voyelle}) + (1 | \text{loc.:DAF})$

**Référence :** la f0 médiane de la voyelle non accentuée (V1 dans SA initial et V4 dans SA final) des locutrices dans la « condition uniquement DAF » dans les essais sans DAF

Groupe 1 = Groupe Positif ; Groupe 2 = Groupe Négatif

Phrase-SA		Estimée	Écart-type	DL	t-valeur	p-valeur
CV-initial (3472 obs.)	<b>Intercept</b>	217,540	7,351	44,642	29,591	<0,001
CVC-initial (3486 obs.)	<b>Intercept</b>	214,850	6,967	44,700	30,838	<0,001
CCV-initial (3506 obs.)	<b>Intercept</b>	213,600	7,113	44,672	30,029	<0,001
CV-final (2333 obs.)	<b>Intercept</b>	207,240	7,864	44,744	26,355	<0,001
	DAF60	0,213	1,489	67,687	0,143	0,887
	DAF120	-0,665	1,491	68,116	-0,446	0,657

Phrase-SA		Estimée	Écart-type	DL	t-valeur	p-valeur
	DAF60:V5	-3,375	1,619	2227,301	-2,085	<b>0,037</b>
	DAF120:V5	-7,209	1,622	2227,085	-4,446	<b>&lt;0,001</b>
	Groupe1:DAF60:V5	5,228	2,730	2226,901	1,915	0,056
	Groupe2:DAF60:V5	4,550	2,866	2226,881	1,588	0,112
	<b>Groupe1:DAF120:V5</b>	<b>7,591</b>	<b>2,747</b>	<b>2226,954</b>	<b>2,763</b>	<b>0,006</b>
	<b>Groupe2:DAF120:V5</b>	<b>6,505</b>	<b>2,898</b>	<b>2226,848</b>	<b>2,245</b>	<b>0,025</b>
	<b>Intercept</b>	206,463	8,264	44,679	24,982	<b>&lt;0,001</b>
CVC-final (2357 obs.)	DAF60	-0,304	1,414	87,906	-0,215	0,830
	DAF120	-1,547	1,423	89,698	-1,087	0,280
	DAF60:V5	-2,442	1,844	2256,096	-1,324	0,186
	DAF120:V5	-4,100	1,855	2254,966	-2,211	<b>0,027</b>
	<b>Intercept</b>	205,800	7,452	45,150	27,612	<b>&lt;0,001</b>
CCV-final (2340 obs.)	DAF60	1,335	1,618	85,070	0,825	0,411
	DAF120	0,483	1,620	85,810	0,298	0,766
	DAF60:V5	-5,782	2,002	2241,000	-2,888	<b>0,004</b>
	DAF120:V5	-8,715	2,003	2238,000	-4,351	<b>&lt;0,001</b>
	<b>Groupe1:DAF120:V5</b>	<b>7,809</b>	<b>3,381</b>	<b>2237,000</b>	<b>2,310</b>	<b>0,021</b>

Tableau 16 Résumé des résultats des six modèles linéaires mixtes comparant l'effet du DAF sur l'intensité médiane des voyelles dans chaque SA (initial ou final) entre les locutrices dans la « condition uniquement DAF » et les locutrices dans la « condition perturbations multiples » dans l'expérience II (cf. chapitre VI 2.3.2)

**Modèle :** Intensité ~ Groupe + DAF + Voyelle + Groupe:DAF + Groupe:Voyelle + DAF:Voyelle + Groupe:DAF:Voyelle + (1 | loc.:Voyelle) + (1 | loc.:DAF)

**Référence :** l'intensité médiane de la voyelle non accentuée (V1 dans SA initial et V4 dans SA final) des locutrices dans la « condition uniquement DAF » dans les essais sans DAF

Groupe 1 = Groupe Positif ; Groupe 2 = Groupe Négatif

Phrase-SA		Estimée	Écart-type	DL	t-valeur	p-valeur
	<b>Intercept</b>	0,908	0,012	46,940	77,209	<b>&lt;0,001</b>
CV-initial (3472 obs.)	DAF60	-0,009	0,004	66,440	-2,086	<b>0,041</b>
	DAF120	-0,017	0,004	67,610	-4,109	<b>&lt;0,001</b>
	<b>Groupe2:DAF120</b>	<b>0,014</b>	<b>0,007</b>	<b>90,420</b>	<b>2,046</b>	<b>0,044</b>
	<b>Intercept</b>	0,926	0,009	48,640	99,758	<b>&lt;0,001</b>
CVC-initial (3486 obs.)	V2	-0,045	0,013	44,420	-3,500	<b>0,001</b>
	DAF60	-0,002	0,004	60,550	-0,596	0,553
	DAF60:V2	0,008	0,004	3378,000	2,104	<b>0,035</b>
	DAF120	-0,009	0,004	63,620	-2,329	<b>0,023</b>
	<b>Groupe2:DAF120</b>	<b>0,018</b>	<b>0,007</b>	<b>81,700</b>	<b>2,755</b>	<b>0,007</b>
	<b>Intercept</b>	0,924	0,007	49,040	126,009	<b>&lt;0,001</b>
CCV- initial (3508 obs.)	Groupe1	0,031	0,011	50,380	2,827	<b>0,007</b>
	Groupe2	0,030	0,012	50,620	2,575	<b>0,013</b>
	DAF60	0,002	0,003	63,650	0,677	0,501
	DAF120	-0,006	0,003	64,500	-1,788	0,078
	DAF60:V2	0,007	0,003	3406,000	1,995	<b>0,046</b>
	DAF120:V2	0,009	0,003	3406,000	2,813	<b>0,005</b>
	<b>Groupe2:DAF120:V2</b>	<b>-0,012</b>	<b>0,006</b>	<b>3406,000</b>	<b>-2,050</b>	<b>0,040</b>
CV-final	<b>Intercept</b>	0,924	0,012	46,770	75,066	<b>&lt;0,001</b>

Phrase-SA		Estimée	Écart-type	DL	t-valeur	p-valeur
(2338 obs.)	V5	-0,081	0,017	44,200	-4,711	<0,001
	DAF60	0,004	0,004	69,160	0,873	0,386
	DAF120	-0,005	0,004	69,560	-1,171	0,246
	DAF60:V5	-0,010	0,004	2236,000	-2,248	<b>0,025</b>
	DAF120:V5	-0,017	0,004	2235,000	-3,974	<0,001
	Groupe1:DAF60:V5	0,014	0,007	2236,000	1,916	0,056
	Groupe2:DAF60:V5	0,014	0,008	2236,000	1,825	0,068
	<b>Groupe1:DAF120:V5</b>	<b>0,019</b>	<b>0,007</b>	<b>2236,000</b>	<b>2,677</b>	<b>0,007</b>
	<b>Groupe2:DAF120:V5</b>	<b>0,018</b>	<b>0,008</b>	<b>2236,000</b>	<b>2,378</b>	<b>0,018</b>
CVC- final (2366 obs.)	<b>Intercept</b>	0,906	0,014	46,490	65,051	<0,001
CCV- final (2354 obs.)	<b>Intercept</b>	0,902	0,015	46,160	61,501	<0,001
	DAF60	0,010	0,004	72,570	2,401	<b>0,019</b>
	DAF120	0,001	0,004	73,020	0,354	0,724

---

# INTERACTION ENTRE STRUCTURE RYTHMIQUE ET SENS D'AGENTIVITÉ EN PRODUCTION DE LA PAROLE

Pour s'adapter aux imprévus lors de la production de la parole, le système moteur intègre les informations sensorielles (ex. le retour auditif), et bénéficie du groupement rythmique, qui est caractérisé par la prosodie. Cependant, le système sensorimoteur d'un locuteur traite différemment les événements acoustiques relatifs à sa propre voix par rapport à ceux d'autrui. Cette thèse vise à examiner la flexibilité de la production de la parole en analysant le rôle organisateur de la prosodie et la sensation subjective d'un locuteur sur le contrôle de sa voix (c.-à-d. le sens d'agentivité relatif à sa voix).

Des expériences des perturbations du retour auditif ont été menées chez des locutrices francophones. Avec le retour auditif retardé (DAF), la différence de durée entre les voyelles accentuées et non accentuées s'est accrue, soulignant une plus grande flexibilité durant la production des accents. De plus, le DAF a induit une réorganisation du rythme de la parole avec un renforcement du groupement syllabique. Avec un décalage continu de la fréquence fondamentale ( $f_0$ ) du retour auditif, la majorité des locutrices ont aligné leur  $f_0$  sur celle modifiée du retour auditif, suggérant que leur système sensorimoteur aurait traité la voix perçue comme une entrée externe. La présence simultanée du DAF et d'un décalage de la  $f_0$  a entraîné une réduction des effets du DAF par rapport à la condition sans décalage de la  $f_0$ . Cette observation suggère une réduction du sens d'agentivité relatif à la voix chez les locutrices, ainsi qu'une interaction entre l'organisation rythmique et le sens d'agentivité dans les processus sensorimoteurs de production de la parole.

**Mots-Clés :** contrôle sensorimoteur de la parole, retour auditif, rythme de la parole, structure prosodique, sens d'agentivité, alignement phonétique, traitement de la parole

---

# INTERACTION BETWEEN RHYTHMIC STRUCTURE AND SENSE OF AGENCY IN SPEECH PRODUCTION

To adapt to unforeseen circumstances during speech production, the motor system integrates sensory information (e.g., auditory feedback) and benefits from rhythmic grouping, which is characterized by prosody. However, a speaker's sensorimotor system processes acoustic events related to their own voice differently from those of others. This thesis aims to examine the flexibility of speech production by analyzing the organizing role of both prosody and a speaker's subjective sensation of control over his voice (i.e., the sense of agency related to his voice).

Experiments of auditory feedback perturbations were conducted with French-speaking female speakers. With delayed auditory feedback (DAF), the duration difference between accented and unaccented vowels increased, highlighting greater flexibility during accent production. Furthermore, DAF induced a reorganization of speech rhythm with enhanced syllabic grouping. With a constant shift in the fundamental frequency ( $f_0$ ) of auditory feedback, the majority of female speakers aligned their  $f_0$  with the modified auditory feedback, suggesting that their sensorimotor system processed the perceived voice as an external input. The simultaneous presence of DAF and an  $f_0$  shift resulted in a reduction of DAF effects compared to the condition without an  $f_0$  shift. This observation suggests a reduction in the sense of agency related to the voice among female speakers, as well as an interaction between rhythmic organization and sense of agency in sensorimotor processes of speech production.

**Key words:** sensorimotor control of speech, auditory feedback, speech rhythm, prosodic structure, sense of agency, phonetic alignment, speech processing